AAMRL-TR-89-023

AD-A215 405

SUBJECTIVE WORKLOAD ASSESSMENT TECHNIQUE (SWAT): A USER'S GUIDE (U)

Gary B. Reid

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

Scott S. Potter
Jeine R. Bressler

SYSTEMS RESEARCH LABORATORIES, INC.

JULY 1989

INTERIM REPORT FOR PERIOD JUNE 1986 - OCTOBER 1988

Approved for public release; distribution is unlimited.

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-89-023

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-89-023	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Armstrong Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable) AAMRL/HEG	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Wright-Patterson Air Force Base, Ohio 45433	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-85-C-0541
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62202F PROJECT NO. 7184 TASK NO. 14 WORK UNIT ACCESSION NO. 07

11. TITLE (Include Security Classification)
Subjective Workload Assessment Technique (SWAT): A User's Guide (U)

12. PERSONAL AUTHOR(S)
Reid, Gary B., Potter, Scott S*, Bressler, Jeine R*

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 6/86 TO 10/88	1989 July	115

16. SUPPLEMENTARY NOTATION
*Systems Research Laboratories, Inc., 2800 Indian Ripple Road, Dayton, Ohio 45440

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
05	08		Mental Workload; Conjoint Measurement; SWAT; Subjective Workload; Conjoint Scaling
12	01		Scale Development; Event Scoring

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report is to serve as a user's guide to accompany software Version 3.1 for the Subjective Workload Assessment Technique (SWAT), a tool for measuring mental workload. The guide has been developed as a "how-to" manual for implementing SWAT. Specifically, this report begins with a general overview of workload, describes in detail all aspects of the scale development phase, provides general information relating to the event scoring phase, and finally goes into a step-by-step explanation of each manu and screen of the program.

Various topics are covered including descriptions of the three SWAT dimensions, use of conjoint measurement and scaling, card sort procedures and analysis, methods of prototyping, event scoring, and data analysis. Each "how-to" section that the user will need has been provided in an appendix so that the appropriate section can be separated from the rest of the report for handy reference.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Gary B. Reid	22b. TELEPHONE (Include Area Code) 513-255-8749 22c. OFFICE SYMBOL AAMRL/HEG

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

SUMMARY

This user's guide along with its appendices is the source document for Version 3.1 of the Subjective Workload Assessment Technique (SWAT), a procedure for measuring mental workload. This particular version of SWAT was developed to be more accessible in terms of both equipment requirements and user friendliness than the previous versions that required a mainframe computer. The basic equipment needed to implement Version 3.1 includes a microcomputer running the MS DOS operating system, 512K internal memory, and two floppy disks or one hard disk and one floppy disk, along with the user's guide. The program can analyze scale development data for up to 30 subjects.

The guide has been written as a "how-to" manual for the first phase of SWAT, Scale Development, but some explanations concerning the second phase, Event Scoring, are included. The report begins with a general overview of workload and proceeds to various topics encompassing the implementation of this particular subjective workload measurement technique. These include sections on descriptions of the three SWAT dimensions, use of conjoint measurement and scaling, card sort procedures and analysis, methods of prototyping, event scoring, and data analysis. Appendix F is a step-by-step explanation of how to use the analysis program, including system requirements, information for getting started, and a description of each menu and screen. This section and each of the other how-to sections have been provided as appendices so that the appropriate part can be separated from the rest of the report and used independently as needed. As an aid to users who need more information on mental or subjective workload, or studies where SWAT has been used, an extensive bibliography has been included.

PREFACE

This report was written as part of an ongoing effort by the Workload and Ergonomics Branch of the Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL) to facilitate the use and ease of implementation of SWAT in both laboratory and operational settings. It was performed under Work Unit 7184-14-07. The effort is supported by Systems Research Laboratories, Inc. (SRL), Dayton, Ohio, under Contract Number F33615-85-C-0541. Mr. Robert Linhart is the contract monitor.

We wish to acknowledge Dr. Thomas E. Nygren of The Ohio State University, who created the original analysis software for Version 3.1, and Mr. Brian Porter of SRL for his time, effort, and patience in producing the user interface for this version.

The disk which this report refers to is Version 3.1 of SWAT and contains four files. They are: (1) SWATPRGM.BAT, (2) MAIN.EXE, (3) SWAT.DAT, and (4) TEST.DAT.

To run the program, type "SWATPRGM" and (RETURN). The first three files which comprise the executable program must be kept together, while the fourth file is a sample data set included for instructional and testing purposes. Since the program needs to write a file to the program disk, it is not write-protected. Therefore, it is advisable to make a backup copy of the disk.

You are asked not to distribute the program disk. We would like to maintain an accurate list of recipients to be able to supply program updates or advise users of problems, should they be discovered. Potential users can obtain the program by writing to:

Mr. Gary B. Reid
AAMRL/HEG
Wright Patterson Air Force Base, Ohio 45433-6573

TABLE OF CONTENTS

Section		Page
1	INTRODUCTION	9
2	SCALE DEVELOPMENT	11
	2.1 INTRODUCTION	11
	2.2 DESCRIPTION OF THE DIMENSIONS	11
	2.2.1 Time Load	12
	2.2.2 Mental Effort Load	12
	2.2.3 Psychological Stress Load	12
	2.3 THE CONJOINT MEASUREMENT MODEL	13
	2.3.1 Introduction	13
	2.3.2 Axiom Tests	15
	2.3.2.1 Independence	16
	2.3.2.2 Joint Independence	16
	2.3.2.3 Double Cancellation	17
	2.3.3 Scale Development Data Collection	17
	2.3.4 Axiom Testing Programs	19
	2.3.5 Scaling	20
	2.3.5.1 MONANOVA	20
	2.3.5.2 Johnson's Algorithm	21
	2.3.5.3 Final Scaling Solution	22
	2.3.5.4 Subscales	23
	2.4 CARD SORT PROCEDURE	23
	2.4.1 Introduction	23
	2.4.2 Timing of Card Sorts	24
	2.4.3 Additional Comments	25
	2.5 CARD SORT ANALYSIS	27
	2.5.1 Introduction	27
	2.5.2 Prototyping	27
	2.5.3 Sample Data Analysis	30
3	EVENT SCORING	50
	3.1 INTRODUCTION	50
	3.2 TASK DEFINITION	51
	3.2.1 Laboratory Tasks	51
	3.2.2 Operational Evaluations	51
	3.3 DATA ANALYSIS	53

TABLE OF CONTENTS (continued)

Section		Page
4	SUMMARY	55
Appendix A	CARD DECK FACSIMILE	56
Appendix B	OUTLINE FOR VERBAL CARD SORT INSTRUCTIONS: COLLEGE STUDENTS	62
Appendix C	OUTLINE FOR VERBAL CARD SORT INSTRUCTIONS: PILOT POPULATION	71
Appendix D	WRITTEN CARD SORT INSTRUCTIONS	78
Appendix E	SAMPLE CARD SORT DATA SHEET	83
Appendix F	INFORMATION FOR PROGRAM OPERATION	85
Appendix G	EVENT SCORING INFORMATION	105
	BIBLIOGRAPHY	107
	REFERENCES	114

LIST OF FIGURES

Number		Page
1	Comments and Main Menu for Sample Data Set Indicating Utilization of an Existing File	31
2	Comments and Main Menu Displaying Previously Saved Information	31
3	Data Entry Screen for Sample Data Set	33
4	Printout of Entire Data Set	34
5	Prototype Analysis of All 12 Subjects Indicating the Need for Prototyped Solution	35
6	Summary of Axiom Violations for the Six Time Prototype Subjects	36
7	Summary of Axiom Violations for the Two Effort Prototype Subjects	37
8	Summary of Axiom Violations for the Four Stress Prototype Subjects	38
9	Summary of Axiom Violations for Subject No. 8--Individually	40
10	Summary of Axiom Violations for Subject No. 12--Individually	41
11	Prototype Analysis Screen Indicating Procedure for Excluding Subject No. 12	42
12	Prototype Analysis Screen Indicating Recalculation of the Kendall's Coefficient Excluding Subject No. 12	43
13	Prototype Analysis Screen Indicating Recalculation of the Kendall's Coefficient Including Only the Time Prototype Subjects	44
14	Prototype Analysis Screen Indicating Recalculation of the Kendall's Coefficient for the Stress Prototype Subjects and Subject No. 8	45
15	Summary of Axiom Violations for the Modified Stress Prototype Group	46
16	Scaling Information for the Six Time Prototype Subjects	48
17	Scaling Information for the Five Stress Prototype Subjects	49
18	SWAT Rating Scales	67
19	Event Scoring	68
20	Scale Development	73
21	Equipment Specification Screen	87
22	Comments Section of Main Menu Screen	88
23	Main Menu Screen with Sample Comments	89

LIST OF FIGURES (continued)

Number		Page
24	Data Entry Screen	92
25	Program Setup Screen	94
26	Prototype Analysis Screen	97
27	Axiom Tests Summary Screen	99
28	Scaling Information Screen	101
29	Plot of Rescaled Versus Raw Data Screen	102
30	Scaling Solution Screen	103

LIST OF TABLES

Number		Page
1	RANK ORDERING OF THE SIX PROTOTYPE GROUPS AS IDENTIFIED BY THE RANDOMLY ASSIGNED LETTERS ON THE SWAT CARD DECK	29
2	PROTOTYPE ANALYSIS OF ONE SUBJECT'S DATA	29
3	EXAMPLE OF HOW TO ASSIGN A RESCALED VALUE TO A SUBJECT'S GIVEN RATING	54

Section 1

INTRODUCTION

The Subjective Workload Assessment Technique (SWAT) has been developed in response to a need for a workload measure with known metric properties that is useful in operational or "real-world" environments. Maximum effort has been expended to keep the SWAT data collection as unintrusive as possible. The principal way this has been accomplished is through the application of a scaling procedure known as conjoint scaling. This approach allows responses to be made in the operational setting using only three simple descriptors for each of three factors that have been used to operationally define workload. This approach also minimizes the amount of time required to make responses by keeping down the number and complexity of descriptors that an operator must memorize.

SWAT is divided into two distinct phases: Scale Development and Event Scoring. The Scale Development phase is used to train the subjects on the use of the descriptors and to obtain data concerning how these dimensions combine to create each individual's personal impression of workload. The Event Scoring phase is the experiment or test situation where the investigator is interested in obtaining information about the workload associated with task performance.

This SWAT User's Guide is to accompany Version 3.1 of the software used in the scale development phase of SWAT and is intended to be a "how-to" manual. Discussion of technical issues and research related to the application of SWAT will be a limited part of this manual. A SWAT bibliography is included for readers interested in a more in-depth treatment of these subjects. Before we begin, there is one additional point that needs to be made. As is well known, there are at least two types of workload--mental workload and physical workload. There is an extensive history of physical measurement and physiological measurement that addresses physical workload. We have chosen to use SWAT as a measure of mental workload and, in general, omit consideration of physical workload. Some physical components that are within normal tolerance ranges but provide a source of irritation may be considered and this will be explained in a later section. In the event that a large physical component of workload is expected to be present in a study, the experimenter should plan to obtain appropriate measures to reflect it. Despite the fact that SWAT is a measure of mental workload only, throughout this manual references to mental workload will be shortened to workload. This is done solely for convenience and should not contribute to overlooking other important components of workload.

This manual is organized around an explanation of the two phases of SWAT--the Scale Development phase and the Event Scoring phase. The "how-to" portions have been placed in appendices

to facilitate easy access after you have become reasonably familiar with the SWAT analysis process. You may wish to separate certain appendices from the main document so they may be used as a master for copies or as a quick reference guide for running the SWAT software.

Section 2
SCALE DEVELOPMENT

2.1. INTRODUCTION

The Scale Development phase is the principal aspect which differentiates SWAT from other subjective workload approaches. Usually, descriptors are provided in order to define some number of workload levels (seven, for example), and subjects are carefully trained to know what is represented by each level of the scale. In SWAT, descriptors of components of workload are provided, but the task of the subject is not to learn what the various levels mean but rather to make judgments that allow the investigator to determine how the factors combine for the particular subjects involved in the investigation.

The first requirement in the development of this scaling approach is to establish an operational definition of mental workload. While researchers have not arrived at a consensus about a technical definition of workload, there is considerable agreement that mental workload is a combination of several factors related to task demands, operator state, and time factors. Therefore, workload has been defined for SWAT to be composed primarily of Time Load, Mental Effort Load, and Psychological Stress Load (Reid and Nygren, 1988). Time Load refers to the total amount of time available to an operator to accomplish a task as well as overlap of tasks or parts of tasks; Mental Effort Load is the amount of attention or concentration that is required to perform a task; and Psychological Stress Load is the presence of confusion, frustration, and/or anxiety associated with task performance. This definition is not intended to represent a sufficient technical definition of mental workload; rather, it provides a useful operational definition. It does appear to reflect most of what a majority of people are talking about when they refer to mental workload.

2.2. DESCRIPTION OF THE DIMENSIONS

The three factors (or dimensions) used to operationally define workload have each been further defined by a set of descriptors that specify three levels of each of the dimensions. These dimensions are based largely on the theoretical work of Sheridan and Simpson (1979) in defining pilot workload. We have attempted to generalize the wording of the descriptors in order to create a scale that is applicable to most work situations where mental workload is an anticipated problem. The dimensions are defined in the following paragraphs.

2.2.1. Time Load

The Time Load dimension depends on the availability of spare time and the overlap of task activities. This is closely associated with the use of time line analysis as a primary method of evaluating whether or not a person should be able to accomplish a task. Time Load may be experienced as the rate that events occur or the speed of a system. The three levels are:

1. Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

2. Occasionally have spare time. Interruptions or overlap among activities occur frequently.

3. Almost never have spare time. Interruptions or overlap among activities are frequent or occur all the time.

2.2.2. Mental Effort Load

Mental Effort Load is an indicator of the amount of attention or mental demands that are required to accomplish a task, independent of the number of subtasks or time limitations. With low Mental Effort Load, the concentration and attention required by a task are minimal and thus performance is almost automatic. As Mental Effort Load increases, so does the amount of concentration and attention required. Generally, this is due to the complexity of the task or the amount of information which must be processed by the operator in order to perform adequately.

High demand for mental effort requires total attention or concentration due to task complexity or the amount of information that must be processed. Activities such as performing calculations, making decisions, remembering or storing information, and problem solving are all examples of mental effort. The exact descriptors used are:

1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

3. Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

2.2.3. Psychological Stress Load

Psychological Stress Load refers to conditions that produce confusion, frustration, and/or anxiety during task performance and, therefore, make task accomplishment seem more difficult. At low levels of stress, one feels relatively relaxed. As stress increases, distraction from relevant aspects of the task is caused by factors within the environment or the individual. These factors include such things as motivation, fatigue, fear, skill level, or temperature, noise, vibration, and comfort. Many of these factors can directly affect task performance when they reach high levels. However, for the purposes of SWAT and the measurement of mental workload, we are talking about these factors when they are at relatively low levels but create enough of an irritant that individuals must draw on resources in order to prevent interference with task performance. The specific levels for the Psychological Stress Load dimension are:

1. Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

2. Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

3. High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

2.3. THE CONJOINT MEASUREMENT MODEL

2.3.1. Introduction

There are many composition rules that describe how complex multifactor or multidimensional judgments are formed. One simple yet psychologically useful rule is the additive rule which suggests that independent variables combine in an independent additive fashion to produce an overall joint psychological effect. This additive rule is the model that underlies SWAT. For example, let T1 be a level of the Time factor, E1 be a level of the Effort factor, and S1 be a level of the Stress factor. We might hypothesize that the joint effects of these three factors (which we call workload) could be described as:

$$f(T1, E1, S1) = f1(T1) + f2(E1) + f3(S1) \quad (1)$$

where f, f1, f2, and f3 are separate and identifiable numerical functions (Krantz and Tversky, 1971).

Additive models like the three-factor model illustrated in equation (1) have been and continue to be an important part of many psychological theories. Until recently, however, even for this simple model, there has not been a satisfactory means by which one could simultaneously estimate all four functions--f1, f2, f3, and f. Conjoint measurement theory provides a means to accomplish this objective. Just as important, however, is the aspect of the theory which indicates that only ordinal relations are required among the data points to produce resultant scales that possess interval properties. The implications of this result will become more apparent following the presentation of an introduction to the basic theory of conjoint measurement.

Prior to an introduction to the mathematical foundations of conjoint measurement, it is useful to define two terms that are generally distinguished in the literature (Emery and Barron, 1979; Green and Rao, 1971; Green and Srinivasan, 1978). First, we define conjoint measurement as the procedure whereby we specify, for a given combination rule, the conditions under which there exist measurement scales for the dependent and independent variables, such that the order of the joint effects of the independent variables in the data is preserved by the numerical composition rule. We then define conjoint analysis (sometimes referred to as numerical conjoint measurement) as the procedure whereby the actual numerical scale values for the joint effects and the levels of the independent variables are obtained. Thus, there are effectively two separate and independent processes in the conjoint measurement methodology and, hence, two such processes in SWAT. First, one attempts to find the appropriate combination rule and then, assuming the rule is valid, finds numerical functions that "best" fit the observed order of the joint effects in the data while conforming to the specified rule.

The general theory as outlined by Krantz and Tversky (1971) provides for a series of axioms which, when tested with a set of data, aid in discriminating among four simple polynomial models to determine which of them best fit the set of data. For example, let f1(T1), f2(E1), and f3(S1) represent the subjective scale values associated with these levels for a given individual. We could postulate that the levels of the three factors combine to form an overall judged value for perceived workload, f(T1, E1, S1), via either:

an additive model (as previously described), if

$$f(T1, E1, S1) = f1(T1) + f2(E1) + f3(S1), \qquad (2)$$

a multiplicative model, if

$$f(T1, E1, S1) = f1(T1) * f2(E1) * f3(S1), \tag{3}$$

a distributive model, if

$$f(T1, E1, S1) = f1(T1) * [f2(E1) + f3(S1)], \tag{4}$$

or a dual-distributive model, if

$$f(T1, E1, S1) = f1(T1) + [f2(E1) * f3(S1)]. \tag{5}$$

Note that in the latter three models, the overall value of the combined effect of the three factors, f(T1, E1, S1), could be completely erased if one of the multiplicative factors has a zero level. In this case, it would not matter what the levels of the other factors were. For an additive model, of course, this is not the case, since a zero level of a factor would make only that factor irrelevant for the combined stimulus effect. Since in this and many other applications one would not expect to find a multiplicative factor with this zero level property, most theoretical and empirical research in conjoint measurement has focused on the additive model.

2.3.2. Axiom Tests

The Krantz and Tversky (1971) axioms define five ordinal properties that are useful in differentiating among the models in equations (2) through (5). In addition, all are necessary although not sufficient for the additive model. These are single factor independence, joint factor independence, double cancellation, distributive cancellation, and dual-distributive cancellation. It is clear from the results of a recent Monte Carlo study (Nygren, 1985) that the critical axioms that are used to assess additivity are independence, joint independence, and double cancellation. Hence, these axioms are used in the SWAT analysis to determine if an additive model exists in the data. These three axioms are summarized below.

Distributive cancellation, which is similar to double cancellation but with four antecedent conditions, was found to be extremely weak as a diagnostic tool in determining additivity. In fact, it was found that even for random nonadditive data, the property will be satisfied, on the average, about 75 percent of the time. Dual-distributive cancellation is a very complex property, requiring five antecedent conditions from a 5 x 5 x 5 design to be met in order for the test to even be

possible. It, like distributive cancellation, is not able to reject additivity. Therefore, these two axioms are not performed in this application of conjoint measurement. For a more detailed description and explanation of these properties, the interested reader is referred to Nygren (1982, 1985).

2.3.2.1. Independence

We begin with the fundamental property of independence which can be checked separately for each of the three factors. In general, we say that:

A is independent of B and C whenever

$$(a1,b1,c1) > (a2,b1,c1) \text{ if and only if } (a1,b2,c2) > (a2,b2,c2) \qquad (6)$$

where A, B, and C represent the three dimensions, and a1, a2, and a3 represent three levels within the first dimension. Similarly, b1, b2, and b3 represent the three levels of the second dimension; likewise, c1, c2, and c3 for the third dimension. Thus, independence of A asserts that if a2 > a1 for some combination of levels of factors B and C, then this relation will hold for any other combination of levels of B and C. For SWAT, this axiom can be translated as stating that the ordering of any two levels of the time load dimension, for example, will remain consistent for all combinations of the other two dimensions. Note that this "independence" is really a monotonicity property and not a statistical independence property. Hence, theoretically it would be quite possible, for example, to find Time Load independent (monotonic) of Effort and Stress, but not Effort independent of Stress and Time or Stress independent of Effort and Time. To the extent that the monotonicity or simple independence property holds, we have support for an additive model. Every test of independence of A with B and C requires a 2 x 2 x 2 matrix with two levels of factor A and two combinations of B x C. Thus the total number of possible tests of the property in this case would be 108 for the 3 x 3 x 3 design.

2.3.2.2. Joint Independence

A second form of independence can also be examined in our three factor model. The property known as joint independence states that:

A and B are jointly independent of C whenever

$$(a1,b1,c1) > (a2,b2,c1) \text{ if and only if } (a1,b1,c2) > (a2,b2,c2). \qquad (7)$$

Joint independence of A and B with respect to C indicates that if one combination of A and B is greater than another at a fixed level of C [i.e., (a1,b1) > (a2,b2) at c1], then the ordering should be preserved for any other level of the third factor (c2). If joint independence holds for all pairs of factors, then this implies that independence holds for a single factor. However, the converse is not necessarily true. If simple independence holds for all factors, this does not imply that joint independence will be satisfied for all pairs of factors. We can, of course, state two other forms of the joint independence property for A and C of B, and B and C of A. For SWAT, this asserts that the ordering of any combination of two dimensions (Time Load and Mental Effort Load, for example) will hold for all levels of the third dimension, Psychological Stress Load.

2.3.2.3. Double Cancellation

The third property examined by Krantz and Tversky (1971) is usually referred to as double cancellation stated for factors A and B as:

If (a2,b3,c1) > (a1,b2,c1) and (a3,b2,c1) > (a2,b1,c1), then (a3,b3,c1) > (a1,b1,c1). (8)

Note that double cancellation requires at least three levels of each of the factors A and B, and deals with only two such factors at a time. Hence, it must be satisfied for all pairs of factors. If factors A and B each have three levels, then there will be one possible test of double cancellation for these two factors. For the purposes of SWAT, this axiom tests the consistency in the row, column, and diagonal relations for two factors at a time, holding the third factor constant. Therefore, there are three possible tests of double cancellation for each factor. Note that in this test there are two antecedent conditions, both of which must be met for the test to be performed.

2.3.3. Scale Development Data Collection

Note that these properties are stated in terms of order relations, and only these order relations are necessary to adequately test these properties. Thus, it is sufficient to require each subject to merely present rank-order judgments for each of the stimulus combinations generated by combining levels of the factors. In SWAT, the method for obtaining this information is the card sort. During the scale development phase, each subject is given a deck of 27 cards. A facsimile of these cards has been included in Appendix A. Each of the cards has three statements on it representing one of the possible combinations of the levels of Time Load, Mental Effort Load, and Psychological Stress Load. Therefore, there is a total of 27 possible combinations of the descriptors. Subjects are

asked to sort the cards so that the 27 combinations are rank ordered to reflect the degree of subjective workload, from lowest to highest, represented by each combination.

A card deck facsimile (Appendix A) has been included as both an example and a master to be used in making up your own card decks. The letter printed on each card actually should be placed on the back of the cards when they are printed. Typed capital letters are easier to distinguish than lower case letters. If lower case letters are used, a line should be placed under each letter to aid in discriminating those letters which might be confused if they are turned upside down (e.g. an upside down "u" frequently looks like an "n").

Each letter has been specifically assigned to each card and that letter has a specific workload combination assigned to it. For example, card "N" has a 1-1-1 workload combination and card "K" has a 2-3-1 combination. The 1-1-1 card is composed of the lowest level of Time Load, Mental Effort Load, and Psychological Stress Load, while the 2-3-1 card is made up of the middle level of Time Load, the highest level of Mental Effort Load, and the lowest level of Psychological Stress Load dimension. Examples are:

	Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.
CARD N	Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.
	Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

	Occasionally have spare time. Interruptions or overlap among activities occur frequently.
CARD K	Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.
	Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

NOTE: It is important to keep each letter with the combinations as they appear in Appendix A. The SWAT analysis program is designed to accept a specific order of those combinations when data are input. Changing the letters will affect the input order and, therefore, the outcome of the SWAT analysis.

2.3.4. Axiom Testing Programs

As discussed earlier, in most applications of conjoint measurement methodology, it is the additive representation that is of interest. However, even for an additive model as small as the 3 x 3 x 3 design used in the SWAT methodology, both the testing procedures for the properties mentioned above and the actual scaling procedure for obtaining the numerical scale values become extremely impractical without the aid of a computer based algorithm. Indeed, the actual counting of the number of tests, the number of successful orderings, and the number of violations of the three axioms described above can become computationally overwhelming very quickly. The SWAT program was designed to meet these computational axiom testing needs practically and efficiently. There are several programs which have been designed to do the testing of the preceding axioms. The SWAT program is a combination of several of these programs.

One attempt to develop a general diagnostic program for testing the conjoint measurement axioms was made by Holt and Wallsten (1974). Their program, CONJOINT, was designed to test each of the axioms mentioned above. CONJOINT was written in PL/1 and has been modified to run on an IBM 370 or Amdahl 470 operating system. Ullrich and Cummins (1973) developed two other programs, PCJM and PCJM2, written in FORTRAN, to do essentially the same thing as CONJOINT. There are, however, several differences between the programs which make both useful as diagnostic tools. For a more detailed description of these programs, refer to Nygren (1982).

SWAT is a combination of what is believed to be the most useful parts of these programs. First, SWAT provides some of the same information as CONJOINT for testing conjoint measurement axioms. However, SWAT also provides a more detailed analysis of violations of these axioms, especially for the critical axioms of simple independence and joint independence. In addition, SWAT is written in FORTRAN, whereas CONJOINT is written in PL/1. SWAT employs some of the same algorithms used in PCJM for examining the axioms. SWAT, however, makes some very important corrections to logical and theoretical errors made by the PCJM analysis of the conjoint measurement axioms.

2.3.5. Scaling

After the axiom tests have been completed and an additive model has been found to be an adequate representation of the data, the SWAT program calculates a scaling solution for the data. By a scaling solution we mean that numerical values can be found for each of the levels of the three factors and their additive combinations that will maintain the order of the subjects' card sort and conform to an additive model.

A number of algorithms are now available for obtaining scaling solutions for multiplicative, distributive, and dual distributive models as well as an additive model. Two of these nonmetric algorithms, MONANOVA (Kruskal, 1965) and NONMETRG (Johnson, 1973), are used for SWAT to provide a best-fitting scaling solution to represent the card sort data. The definition of "best fitting" is what differentiates the two procedures, as will be described below. Being nonmetric scaling procedures, both algorithms attempt to determine the best fitting set of interval-scaled values for the levels of the dimensions and their combined effect based only on rank ordering of the combinations of the dimensions. Therefore, the SWAT procedure begins by rank ordering the data from the smallest to the largest, if they are not already in that form. From this point on, only this rank order of the data and not the data values themselves are used in the analysis.

2.3.5.1. MONANOVA

The first approach, MONANOVA, finds and applies a monotonic transformation to the original card sort data such that a set of distances can be estimated for the combinations of levels with the constraint that the distances fit the transformation in a least squares analysis. To begin, an arbitrary set of initial scale values for the levels of the factors is formed to produce initial estimates of the 27 stimulus combinations. From these initial scale values, a matrix of what is called disparities is formed. Disparities are transformed values that are monotonic with the original data and as close as possible to the initial set of workload scale values. Next, a badness-of-fit measure, STRESS, is computed to determine how closely the monotonically transformed disparity values match the estimated scale values from the additive model. STRESS is formed by finding the square root of the sum of the squared deviations between the disparity values and the estimated stimulus values.

If the original rank data are in perfect agreement with an additive representation, then monotonically transformed disparities will be found that, when suitably normalized, are identical to the estimated stimulus scale values, producing a STRESS value of zero. Subjects' data are, however, generally not without some random error. Typically then, the algorithm will not find a STRESS value of zero. In these cases, the algorithm works iteratively. Following the computation of

STRESS, the estimated stimulus scale values are recalculated via a least-squares estimation procedure similar to that employed in standard regression analysis. The partial derivative of STRESS with respect to each scale value is found, and a numerical analysis procedure known as the method of gradients is used to find a new set of best-fitting (in the least squares sense) stimulus scale values. New disparities are formed, a new STRESS value is computed, and the iterative process is continued until no improvement in STRESS can be found. Following the last iteration, the estimated scale values for the 27 stimulus combinations are found and are normalized so that combination (1,1,1) has a scale value of 0 and (3,3,3) has a scale value of 100.

This approach treats inconsistencies in the arrangement of levels in the card sort as being meaningful information. Therefore, inconsistencies will cause the scaling to incorporate ties in the levels. As a result, this is the preferred algorithm when a subject's conceptualization of workload is defined by a model with fewer than three levels of each of the dimensions. For example, a subject may feel that Time Load does not really make any difference in how high his workload is up to the point of not being able to keep up with the task. From that point on, this subject might think that Time Load is an extremely important factor. In this case, the subject would probably treat the level 1 and level 2 descriptors as equivalent while level 3 would be distinctly different. In other words, the subject would have only two levels of Time Load and, perhaps, three levels of Mental Effort Load and three levels of Psychological Stress Load. Therefore, the SWAT program would automatically select and present the MONANOVA scaling solution as the best fitting solution.

2.3.5.2. Johnson's Algorithm

The second scaling procedure in SWAT, NONMETRG, is used to provide another scaling of the data, this time based on a badness-of-fit measure different from STRESS. This measure, THETA, differs from STRESS in that it is based on a pairwise method in which the differences in scale values for all possible pairs of stimuli (351 pairs for the 27 stimuli in SWAT) are compared with the differences in the original ranks.

As in the previous scaling algorithm, this routine starts by finding a set of estimates of the stimulus scale values. For efficiency, it uses the final estimates found by the STRESS-based procedure as these starting values. If the data do, in fact, perfectly conform to an additive model, the procedure stops after one iteration, since the scale values have already been determined. If the data are not perfectly additive (as is usually the case), then the badness-of-fit measure THETA is computed by summing the differences in scale values for all pairs of stimuli for which the original ranks are not in the same order as the estimated scale values. This sum is then normalized by dividing by the sum of all differences in scale values and taking the square root. The numerator of this term and,

thus, THETA will be zero if all pairs of ranks and pairs of estimated scale values are in the same order. As in the case of STRESS, the partial derivative of THETA (actually THETA-squared) is taken with respect to each scale value in order to find new estimates that will minimize the differences in scale values for which there are incorrect pairwise orderings. The iterative procedure is then continued until no significant improvement in the estimated scale values (i.e., that will minimize THETA) can be found. Following the last iteration (as in MONANOVA), the scale values are normalized to a range of 0 to 100.

This THETA measure is strongly related to Kendall's Tau coefficient, although they are not a simple function of one another. Therefore, Kendall's Tau is included as another goodness-of-fit indicator in the NONMETRG solution. In SWAT for example, for a set of ranks that is perfectly additive, Tau will be 1.0, indicating that all 351 pairs of estimated scale values are in the same order as the 351 pairs of ranks. For data containing a slight error, it is still possible for THETA to be 0.0 (by producing tied scale values) but for Tau not to be equal to 1.0.

It may at first seem redundant to perform two scaling procedures in SWAT, since both will yield identical results for perfectly additive data and generally very similar results. However, the differences in handling of inconsistencies (error) provide a complementary approach to the conjoint scaling problem. MONANOVA interprets inconsistencies in the data as error and the order from the card sort is maintained. The scale values are forced to be different, resulting in a scale that would not properly reflect the two level model in the previous example. On the other hand, NONMETRG will give a better solution when a subject's card sort conforms to the typical model with three levels of each of the three dimensions but has a number of inconsistencies that are, in fact, random error. If the program cannot find a pattern of ties in the ordering that explains inconsistencies, then they are assumed to be error and NONMETRG is the selected scaling solution.

2.3.5.3. Final Scaling Solution

As previously described, SWAT incorporates both MONANOVA and NONMETRG into the scaling solution. MONANOVA is performed first, and produces a set of rescaled values for each of the three levels of the three dimensions. To aid in efficiency, these rescaled values are then used as input for the NONMETRG algorithm. After both solutions have been calculated, the SWAT program chooses the best solution from the two algorithms. This is based on a rule which assumes that a difference of less than six units between the rescaled values of any two levels of a dimension implies essentially equivalent levels. This occurrence is best fitted by the MONANOVA algorithm. If the differences between each pair of all three levels of any given dimension are greater than six, the program will choose the solution from the NONMETRG algorithm. The selected solution

automatically will be displayed by the program. However, both solutions may be examined by selection of the proper options in the program, as described in Appendix F.

2.3.5.4. Subscales

Whichever scaling solution is used, the result is rescaled values with interval-level properties for each of the three levels of the three dimensions. Typically, these are combined to form the 27 combinations of workload levels. However, it should be noted that these subscale values, as we call them, can also be used independently as a possible diagnostic approach to workload assessment. Studies have been performed to investigate the differential sensitivity of the three subscales (Eggemeier, McGhee, and Reid, 1983; Potter and Acton, 1985). Both of these studies indicated potential advantages to analyzing the subscales as well as the overall combined values.

2.4. CARD SORT PROCEDURE

2.4.1. Introduction

The principal reason for completing the card sort is to generate data that are used to produce a scaling solution tailored to the group's or individual's perception of workload. This is one aspect of SWAT which is different from most other subjective workload assessment approaches. The results of the card sort are analyzed by the conjoint scaling program to produce an interval-level workload scale.

Since the results from card sorts are used to generate the workload scale, the card sort session is the key to a successful application of SWAT. The subject must be convinced of the importance of providing the best possible information regarding how he or she perceives (trades off) the three dimensions which we have defined as being the primary contributors to workload. Inaccurate or invalid card sort information can have a considerable effect on the results of the experiment.

Aside from the scale generation, there are several other very important aspects of the card sort which need to be emphasized. Primarily, the card sort procedure serves as training for the subsequent Event Scoring phase of SWAT. After sorting the cards, subjects are very familiar with the use of the three dimensions and their levels. Consequently, only slight additional training will be required (see Section 3 for more information on Event Scoring).

Secondly, the card sort provides motivation for subjects to take the rating scale seriously. One problem inherent in traditional subjective measures is that of gaining subject acceptance of the

rating scale being used. If subjects reject the technique or take the rating task too lightly, then the chances of obtaining accurate ratings are greatly reduced. Our experience has indicated that performing the card sort provides subjects with a feeling of greater involvement and, thus, facilitates seriousness and greater reliability of ratings.

Appendices B, C, and D provide information pertinent to the card sort procedure. Appendix B provides an outline which is to serve as a guide for verbal instructions for a typical sample of college students. Similarly, Appendix C is an outline of instructions geared toward a pilot population. You will notice that these two sets of instructions have some similarities and yet they are not identical. This illustrates the point that you should "tailor" the instructions to fit the specific application and subject population. These samples are included to indicate some of the points which should be included in a briefing, but not necessarily read verbatim to the subjects. In writing a set of instructions, points and approaches from both of these sets, as well as others, may be needed. While these instructions are intended to be given verbally, some investigators may wish to have only a written set for their subjects. In this case, a modification of these appendices may be used in conjunction with Appendix D. Appendix D, or a modification of Appendix D, is a set of instructions which nearly all subjects should read before beginning the card sort. These instructions are designed to reemphasize points made in the briefing (Appendix B or C) and provide additional points/comments which are necessary. Appendix D may be copied and made available to all subjects.

Appendix E is a master card sort data sheet which may be used to record the subjects' card sort data. It also should be copied prior to use. The procedure for recording data is as follows. Take a deck of sorted cards; identify which card represents the lowest workload combination (typically this is card N). Find this letter on the data sheet and assign it a rank of one (1). Take the next card from the deck, find its corresponding letter on the data sheet and assign it a rank of two (2). Continue in this manner until all 27 letters on the data sheet have been assigned a rank. Perform this same operation for each subject's card deck. This matrix of numbers will be the input data for the analysis program. The procedure for entering data into the program will be described in Appendix F, Data Analysis Procedures.

2.4.2. Timing of Sort Cards

It is preferable that the sorts be made prior to Ev[illegible] Scoring (participating in the experiment) because of the training value associated with repeated reading of the descriptors. However, as far as conjoint measurement theory is concerned, the sort can be obtained at any time. There are no known requirements regarding elapsed time between card sorts and Event Scoring. In past

experiments, the same subjects have performed sorts as far apart as a year. Over 80 percent of the subjects produced sorts correlating .90 and above. Less than 5 percent gave sorts whose order reflected different combination rules. Thus, the indication is that, if the person is really trying to give an honest evaluation of how these dimensions combine, his/her judgment remains relatively stable for substantial time periods.

Conversely, requiring subjects to repeat the card sort several times could result in annoying them and cause them to be careless in the ordering simply to complete the task as quickly as possible. The result would be an ordering of the cards that does not faithfully represent the subjects' composition rule and a potentially inconsistent ordering which could not be modeled.

In general, it is recommended that the subjects be scheduled for a 1-hour scale development session, either individually or in groups. This session should be dedicated time free from other time constraints. The experimenter should provide a complete explanation about the card sort and make every effort to obtain a sort that accurately reflects the subjects' opinions regarding the combination of the dimensions. The experimenter can then use the subjects' same card sorts data for more than one experiment or for several sessions of the same experiment. There are, of course, certain experimental questions that may require additional sorts from subjects. The general rule is meant to apply to studies such as system evaluations.

2.4.3. Additional Comments

The following is a list of points which commonly arise concerning SWAT. They are answers to the most commonly asked questions concerning card sort procedures.

1. The card sorts can be obtained in group sessions or individual sessions or some of both as study constraints dictate.

2. A strategy that is suggested to the subject for accomplishing the card sort is to make three stacks of approximately nine cards each. One stack represents relatively low workload, the middle stack represents moderate workload, and the other stack represents relatively high workload. Each stack is then ordered 1 to 9 for low to high workload within a stack. The stacks are then put together and the order rechecked.

3. Subjects should be assured that there is not a correct order to the card sort. What is desired, however, is that the order should represent their opinion as to what constitutes workload to them personally and that it may differ from person to person. They are

instructed, for purposes of the sort, to think of the set of descriptors on each card as representative of all workload events like this. It is important that the subject thinks of events that he/she has experienced so that judgments of the joint effects can be made between the different combinations. Their sort should not be based on the anticipation of the upcoming study or experiment.

4. Subjects are permitted to write on the cards if they desire. However, you should request that they try to place the cards in order based upon the word descriptions rather than by creation of an algorithm based upon the level numbers. Occasionally, the data have the appearance of having been generated by application of an algorithm, and marks on the cards usually support this observation.

5. Talking is permitted during the session, but you should request that they express their own opinion in their ordering rather than attempting to develop a consensus sort.

6. It should be pointed out that the cards have letters on the back which have been randomly assigned. The deck is placed in alphabetical order so that all subjects have the same starting configuration. The subjects should be warned against trying to use the letters as indicators of which is higher or lower.

7. A good conscientious effort usually takes at least 20 minutes and may take as long as an hour. Be suspicious of not convincing the subject of the importance of the sort if he/she finishes too quickly.

8. In some instances, a subject's card sort may contain an unacceptable number of axiom test violations, as described in Section 2.3.2 and to be described in Section 2.5.3. In the event of this occurrence, there are two alternatives. First, if the subject's data exhibit an identifiable pattern (such as Mental Effort Load most important, Time Load second-most important, and Psychological Stress least important) but an unacceptable number of axiom test violations, it is possible that when this subject's data are averaged with others, the group errors will be acceptable due to the strength of multiple subjects. However, if this is not the case, another sort may be needed. It is important that the subject not feel that his/her sort was "wrong," but that since this is a communication process, there was a possible breakdown in the exchange of information (emphasize that you may not have explained the task very well). Then, reiterate the main points and try to elicit questions from the subject before starting the new card sort. It is advisable that the additional sorts be performed one-on-one, rather than in groups.

If the errors are not too numerous or if another sort is not possible due to study constraints, it may be possible to reach an acceptable solution through the use of paired comparisons. In this approach, subjects are presented with pairs of cards which violated the predominant rule in their sort and asked to compare each pair independently by stating which of the two is higher in workload. Typically, these pairs are indicated by violations of the independence axiom. A complete listing of the violations can be obtained from the program (see Appendix F for more information). In this manner, it is possible to assess whether the violations were simply inadvertent or the subject actually was giving contradictory information. If the subject follows a certain pattern, he/she may be placed in the appropriate prototype group. If not, additional instructions and possibly an entire new sort may be necessary.

2.5. CARD SORT ANALYSIS

2.5.1. Introduction

The card sort analysis is performed to accomplish two objectives. First of all, the conjoint measurement algorithm performs the axiom tests to assess the validity of an additive model for the data. Secondly, the scaling algorithms produce interval-level rescaled values for each of the levels of the three dimensions. This is accomplished by using the microcomputer based software. The next two sections describe information concerning prototyping and a sample data analysis section for instructional purposes. Appendix F contains detailed instructions for program operation and is designed to be a reference guide for using the program.

2.5.2. Prototyping

One of the unique aspects of SWAT is what is known as prototyping. Prototyping refers to a procedure of stratifying the subjects into homogeneous groups based on their perceptions of the relative importance of the three dimensions included in SWAT. The Prototype screen displays the results of how the individual subjects prototype. A subject who prototypes "time" considers the Time Load dimension to contribute the heaviest to his perception of workload. A subject who prototypes "stress" would consider the Psychological Stress Load dimension to contribute the heaviest. These are just two examples of the three "main" prototyping groups. The six "possible" groups are listed in the output as TES, TSE, ETS, EST, SET, and STE.

To calculate prototype group membership, every subject's data are correlated against the six different strings of data that represent these respective prototype groups, which are presented in Table 1. The pattern of correlation coefficients of the six groups determines to which group a subject belongs. As shown in Table 2, subject No. 1 correlates the highest with SET and STE and so he is considered a stress subject. In this manner, nearly every subject can be labeled as either a time, effort, or stress subject. A more detailed description of this procedure can be found in Reid, Eggemeier, and Nygren (1982). While a large majority of subjects normally fall into one of the three "main" prototypes, this is not always the case. For some subjects, as with subject No. 4 in Table 2, the two highest correlation coefficients do not both point to a single prototype group. For this subject, the highest correlation is with TES, but the second highest correlation is with ETS. Stress is clearly the least important of the three dimensions, but time is not clearly the most important. In this situation, the subject would be considered to belong to a time/effort prototype group. For this occurrence, the experimenter may wish to create a fourth prototype group and generate a separate scale for this group, or the experimenter may choose to place the T/E subject into whichever one of the three main groups seems to best fit the data. These choices are subjective, and the value of the correlation coefficients and the number of T/E subjects should both be considered.

For simplicity, the remaining discussion will only consider the three main prototypes. At this point, it is important to note the three methods for handling the data and creating the final SWAT scale:

1. Group Scaling Solution: The data from all subjects will be averaged together, and the conjoint scaling algorithm will derive the scale from this average.

2. Prototyped Scaling Solution: Subjects are prototyped as either time, effort, or stress subjects, and each of these three homogeneous groups then has its own SWAT scale. The investigator may override the automatic selection of prototypes and create as many prototype groups or reassign a subject to a prototype as is considered necessary.

3. Individual Scaling Solution: Each subject's data are analyzed separately and a SWAT scale is derived for each subject individually.

The criteria for deriving either a group scaling solution or a prototyped scaling solution is based upon the value of the Kendall's Coefficient of Concordance which is listed immediately before the prototyping section. This coefficient is an index of the degree of intersubject agreement within the card sort. Our experience indicates that a value of approximately .75 and above indicates a relatively homogeneous group of subjects, and only one scaling solution is necessary to capture the

TABLE 1. RANK ORDERING OF THE SIX PROTOTYPE GROUPS AS IDENTIFIED BY THE RANDOMLY ASSIGNED LETTERS ON THE SWAT CARD DECK

	TES	TSE	ETS	EST	SET	STE
N	1	1	1	1	1	1
B	2	4	2	4	10	10
W	3	7	3	7	19	19
F	4	2	10	10	4	2
J	5	5	11	13	13	11
C	6	8	12	16	22	20
X	7	3	19	19	7	3
S	8	6	20	22	16	12
M	9	9	21	25	25	21
U	10	10	4	2	2	4
G	11	13	5	5	11	13
Z	12	16	6	8	20	22
V	13	11	13	11	5	5
Q	14	14	14	14	14	14
ZZ	15	17	15	17	23	23
K	16	12	22	20	8	6
E	17	15	23	23	17	15
R	18	18	24	26	26	24
H	19	19	7	3	3	7
P	20	22	8	6	12	16
D	21	25	9	9	21	25
Y	22	20	16	12	6	8
A	23	23	17	15	15	17
O	24	26	18	18	24	26
L	25	21	25	21	9	9
T	26	24	26	24	18	18
I	27	27	27	27	27	27

TABLE 2. PROTOTYPE ANALYSIS OF SWAT CARD SORT DATA (SPEARMAN RANK CORRELATION, RS, FOR EACH PROTOTYPE)

Subject		TES	TSE	ETS	EST	SET	STE
No. 1	RS	0.62	0.65	0.67	0.72	0.81*	0.79**
No. 2		0.86**	0.88*	0.69	0.65	0.72	0.77
No. 3		0.68	0.80	0.44	0.48	0.84**	0.92*
No. 4		0.81*	0.77	0.80**	0.76	0.67	0.68

*Highest correlation.
**Second highest correlation.

subjects' composite view of workload. A Kendall's Coefficient below .75 usually requires that a separate scaling solution be developed for each of the three main prototype groups, and sometimes for additional groups as previously described. This is needed in order to capture the differential weighting of the three SWAT dimensions that the groups have revealed through their ordering of the descriptor combinations in the card sort procedure.

To determine which method is appropriate for developing the scaling solution, the data must be analyzed using the SWAT program. If the Kendall's Coefficient is .75 or higher, then the scale from that group run may be used with all subjects. If the Kendall's Coefficient is below .75, the subjects should be prototyped as either time, effort, or stress. The appropriate options must be selected on the Program Setup screen in order to produce the required scaling solutions. This procedure will be described with an example in the next section.

2.5.3. Sample Data Analysis

This section will describe the analysis for the sample data set which is included on the SWAT disk. The name of this file is "TEST.DAT." To use this file, it should be copied over to disk drive B before beginning operation. The program disk should be in disk drive A or copied over to a hard disk. This sample data file includes data for 12 subjects that can be used to test the program and as a tutorial on program operation. Throughout this section, emphasis will be on interpreting the analysis. Refer to Appendix F for procedural guidance, if necessary. Appendix F provides detailed information on each of the menus and screens encountered in an analysis. For the novice user, these two sections should probably be used in tandem. It is suggested that these data be entered in another data file for practice in using the program and the results compared with the enclosed figures, since it is essential to become familiar with the data entry procedures. Alternatively, the enclosed file may be retrieved and used for tutorial purposes. In either approach, the results can be compared with the subsequent figures to assure proper program functioning.

To begin operation, type "SWATPRGM." If this is the first time the program has been used, the Equipment Specification screen will appear. After making the appropriate selections for the system being used, the Main Menu will appear. For a new data file, enter the appropriate information on the Main Menu, making sure to enter a file name different than "TEST.DAT." This information will be saved for future reference when the data are saved. To use "TEST.DAT" which exists on disk, bypass all information on the Main Menu screen by entering all RETURN's except for the file name. For this case, enter "TEST.DAT" as the file name. The program will ask whether you want to work with the existing file, to which you should answer "yes." This sequence is presented

in Figure 1. After responding "yes," the information which was previously saved will appear on this screen, as shown in Figure 2.

```
******* COMMENTS AND MAIN MENU *******

TODAY'S DATE:
        (mm/dd/yy)

STUDY NAME:                         FILE NAME:test
        (20 CHARACTERS MAX)               (8 CHARS. MAX)

NUMBER OF SUBJECTS:

COMMENT:
COMMENT:
COMMENT:

                    WARNING
            FILE A:test     EXISTS
            Work with the existing file (y/n) ?


** USE A SEPARATE DATA DISK FOR EACH STUDY **
```

Figure 1. Comments and Main Menu for Sample Data Set Indicating Utilization of an Existing File

```
******* COMMENTS AND MAIN MENU *******

TODAY'S DATE: 02/26/87
        (mm/dd/yy)

STUDY NAME:sample data              FILE NAME:test
        (20 CHARACTERS MAX)               (8 CHARS. MAX)

NUMBER OF SUBJECTS:12

COMMENT:This is a sample data set included with the SWAT program
COMMENT:to demonstrate the procedures for performing conjoint
COMMENT:analysis.  Refer to the User's Guide for instructions.

                    MAIN
                    MENU

    F1  EDIT COMMENTS               F4  EQUIPMENT SPECIFICATION
    F2  DATA ENTRY                  F5  END THE PROGRAM
    F3  PROGRAM SETUP

                MAKE A SELECTION:
```

Figure 2. Comments and Main Menu Displaying Previously Saved Information

In most applications, the first task is to enter data. The Data Entry screen is obtained by entering an F2. When creating a new data set, enter F2 to begin entering data (refer to Appendix F for a complete description of the data entry features). Whether creating a new data file or using "TEST.DAT," after entering data the Data Entry screen should match that of Figure 3. One valuable feature of this program is a data checking algorithm which is invoked as you exit from entering/editing data. To demonstrate this when using "TEST.DAT," enter F2 to enter/edit and then enter an F1 to exit from entering data. As you see, there is a possible error in one of the subjects' data. Typically, it is easier to proof the data by printing it out so that the entire data set can be viewed at once. The entire data set as it should appear by printing it out is presented in Figure 4. Press F3 to print the data and find the error. As a rule for checking data, no matter what order the data are in, every number 1 through 27 should appear only once, and there should be no numbers out of this range. Once you have spotted the error, go back into the data set and make the correction. Then, as after any changes, save the data set.

At this point you can go directly to the Program Setup screen (F4). The first analyses typically performed are the prototype correlations and Kendall's Coefficient of Concordance. The results of these will dictate the direction of further analyses. Choose a "1" on the Program Setup screen and proceed. The results of the prototype correlations and Kendall's coefficient are presented in Figure 5. This screen indicates several important points. First of all, a single group scale would not be appropriate since the Kendall's Coefficient of Concordance is .7380, below the criterion of .75. Secondly, it indicates a potential problem with the effort group. There are only two subjects with highest correlations in the effort group, and their correlations are not very high. Generally a subject with a highest correlation below .80 is cause for further investigation, especially if he/she is in a small group of subjects.

To investigate further, return to the Program Setup screen and choose No. 4, Prototype Axioms, to assess the appropriateness of an additive model on the three groups separately. The results of the axiom tests are presented in Figures 6, 7, and 8. The summary of axiom violations for the time group, as depicted in Figure 6, indicates very good agreement with only a maximum of seven violations for joint independence. In general, the criterion for either independence or joint independence is 20 failures. Enter F1 to go to the next prototype group and display the results for the effort group. As indicated in Figure 7, the correlations for these two subjects do not reach the criterion for acceptance of the analysis. Based on this information, it is advisable to investigate each of these two subjects individually. Before doing that, however, examine the axiom test results for the stress group. As presented in Figure 8, there are no axiom violations for this group.

*** ENTER SUBJECT DATA IN THIS TABLE ***

12 SUBJECTS

F1 SAVE DATA
F2 EDIT/ENTER DATA
F3 PRINT DATA
F4 PROGRAM SETUP
ESC MAIN MENU

CARD	1	2	3	4	5	6	7	8
111 N	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
112 B	7.00	2.00	10.00	2.00	10.00	4.00	10.00	3.00
113 W	10.00	8.00	19.00	3.00	19.00	6.00	19.00	7.00
121 F	5.00	3.00	4.00	6.00	2.00	3.00	2.00	6.00
122 J	12.00	9.00	13.00	5.00	11.00	5.00	11.00	11.00
123 C	9.00	10.00	22.00	4.00	20.00	18.00	20.00	22.00
131 X	6.00	12.00	7.00	7.00	3.00	8.00	3.00	8.00
132 S	8.00	11.00	16.00	8.00	12.00	12.00	12.00	12.00
133 M	21.00	23.00	25.00	9.00	21.00	16.00	23.00	23.00
211 U	2.00	4.00	2.00	10.00	4.00	2.00	4.00	2.00
212 G	3.00	5.00	11.00	11.00	13.00	9.00	13.00	5.00
213 Z	15.00	13.00	20.00	12.00	22.00	20.00	21.00	13.00
221 V	4.00	6.00	5.00	13.00	5.00	7.00	5.00	9.00
222 Q	13.00	14.00	14.00	14.00	14.00	11.00	14.00	20.00
223 ZZ	14.00	15.00	23.00	15.00	23.00	25.00	22.00	14.00
231 K	11.00	16.00	8.00	16.00	6.00	10.00	7.00	10.00
232 E	25.00	17.00	17.00	17.00	15.00	15.00	16.00	21.00
233 R	20.00	24.00	26.00	18.00	24.00	21.00	24.00	26.00

Figure 3. Data Entry Screen for Sample Data Set

12 SUBJECTS

		1	2	3	4	5	6	7	8	9	10
111	N	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
112	B	7.00	2.00	10.00	2.00	10.00	4.00	10.00	3.00	2.00	2.00
113	W	10.00	8.00	19.00	3.00	19.00	6.00	19.00	7.00	3.00	4.00
121	F	5.00	3.00	4.00	6.00	2.00	3.00	2.00	6.00	4.00	3.00
122	J	12.00	9.00	13.00	5.00	11.00	5.00	11.00	11.00	5.00	5.00
123	C	9.00	10.00	22.00	4.00	20.00	18.00	20.00	22.00	8.00	7.00
131	X	6.00	12.00	7.00	7.00	3.00	8.00	3.00	8.00	6.00	6.00
132	S	8.00	11.00	16.00	8.00	12.00	12.00	12.00	12.00	7.00	10.00
133	M	21.00	23.00	25.00	9.00	21.00	16.00	23.00	23.00	9.00	12.00
211	U	2.00	4.00	2.00	10.00	4.00	2.00	4.00	2.00	10.00	8.00
212	G	3.00	5.00	11.00	11.00	13.00	9.00	13.00	5.00	12.00	9.00
213	Z	15.00	13.00	20.00	12.00	22.00	20.00	21.00	13.00	13.00	13.00
221	V	4.00	6.00	5.00	13.00	5.00	7.00	5.00	9.00	11.00	11.00
222	Q	13.00	14.00	14.00	14.00	14.00	11.00	14.00	20.00	14.00	14.00
223	ZZ	14.00	15.00	23.00	15.00	23.00	25.00	22.00	14.00	15.00	18.00
231	K	11.00	16.00	8.00	16.00	6.00	10.00	7.00	10.00	16.00	16.00
232	E	25.00	17.00	17.00	17.00	15.00	15.00	16.00	21.00	17.00	19.00
233	R	20.00	24.00	26.00	18.00	24.00	21.00	24.00	26.00	18.00	22.00
311	H	16.00	7.00	3.00	19.00	7.00	13.00	6.00	4.00	19.00	15.00
312	P	17.00	18.00	12.00	20.00	16.00	17.00	15.00	15.00	20.00	17.00
313	D	19.00	25.00	21.00	21.00	25.00	23.00	25.00	19.00	24.00	23.00
321	Y	23.00	19.00	6.00	22.00	8.00	14.00	8.00	17.00	21.00	20.00
322	A	18.00	20.00	15.00	23.00	17.00	19.00	17.00	16.00	22.00	21.00
323	O	26.00	26.00	24.00	24.00	26.00	26.00	26.00	24.00	25.00	26.00
331	L	22.00	21.00	9.00	25.00	9.00	22.00	9.00	18.00	20.00	24.00
332	T	24.00	22.00	18.00	26.00	18.00	24.00	18.00	25.00	26.00	25.00
333	I	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00	27.00

****

		11	12
111	N	1.00	8.00
112	B	10.00	5.00
113	W	19.00	6.00
121	F	2.00	4.00
122	J	11.00	14.00
123	C	20.00	22.00
131	X	3.00	17.00
132	S	12.00	12.00
133	M	21.00	25.00
211	U	4.00	7.00
212	G	13.00	3.00
213	Z	22.00	16.00
221	V	5.00	2.00
222	Q	14.00	13.00
223	ZZ	23.00	11.00
231	K	6.00	20.00
232	E	15.00	19.00
233	R	24.00	23.00
311	H	7.00	9.00
312	P	16.00	10.00
313	D	25.00	18.00
321	Y	8.00	1.00
322	A	17.00	15.00
323	O	[illegible]	[illegible]
331	L	9.00	24.00
332	T	18.00	[illegible]
333	I	27.00	[illegible]

....

Figure 4. Printout of Entire Data Set

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .7380

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

	SUBJECTS 12						SUGGESTED
SUB. #	TES	TSE	ETS	EST	SET	STE	PROTOTYPE
1	.80	.80	.66	.61	.59	.64	T
2	.80	.79	.74	.71	.68	.70	T
3	.30	.43	.43	.60	1.00	.96	S
4	1.00	.95	.60	.42	.27	.41	T
5	.43	.60	.30	.43	.96	1.00	S
6	.79	.84	.60	.58	.74	.80	T
7	.42	.59	.33	.46	.97	1.00	S
8	.64	.64	.75	.78	.78	.74	E
9	.99	.97	.56	.40	.35	.49	T
10	.97	.94	.68	.56	.48	.58	T
11	.43	.60	.30	.43	.96	1.00	S
12	.45	.42	.70	.75	.67	.58	E

F1 CHANGE PROTOTYPE
F2 PRINT
F3 RETURN TO PROGRAM SETUP
F4 GO TO NEXT OPTION CHOOSEN IN PROGRAM SETUP
ESC MAIN MENU

Figure 5. Prototype Analysis of All 12 Subjects Indicating the Need for Prototyped Solution

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                    PROTOTYPE ANALYSIS
                      TIME PROTOTYPE

          INDEPENDENCE
T INDEPENDENT OF E AND S =   0. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =   0. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =   0. FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF   3  TESTS
DOUBLE CANCELLATION IN E x S =   1. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   3  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =   2. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =   7. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =   2. FAILURES OUT OF 108  TESTS


                    OPTIONS - PROTOTYPES
           F1   GOTO NEXT PROTOYPE
           F2   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
           F3   PRINT SUMMARY OF AXIOM VIOLATIONS
           F4   PRINT COMPLETE AXIOM HISTORY
          ESC   MAIN MENU
```

Figure 6. Summary of Axiom Violations for the Six Time Prototype Subjects

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                PROTOTYPE ANALYSIS
                  EFFORT PROTOTYPE

          INDEPENDENCE
T INDEPENDENT OF E AND S =  28. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =  16. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =  26. FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   1. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN E x S =   1. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   0  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =  23. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =  21. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =  30. FAILURES OUT OF 108  TESTS


                  OPTIONS - PROTOTYPES
       F1   GOTO NEXT PROTOYPE
       F2   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
       F3   PRINT SUMMARY OF AXIOM VIOLATIONS
       F4   PRINT COMPLETE AXIOM HISTORY
      ESC   MAIN MENU
```

Figure 7. Summary of Axiom Violations for the Two Effort Prototype Subjects

******* SUMMARY OF AXIOM VIOLATIONS *******

PROTOTYPE ANALYSIS
STRESS PROTOTYPE

INDEPENDENCE

T INDEPENDENT OF E AND S = 0. FAILURES OUT OF 108 TESTS
E INDEPENDENT OF T AND S = 0. FAILURES OUT OF 108 TESTS
S INDEPENDENT OF T AND E = 0. FAILURES OUT OF 108 TESTS

DOUBLE CANCELLATION

DOUBLE CANCELLATION IN T x E = 0. FAILURES OUT OF 3 TESTS
DOUBLE CANCELLATION IN E x S = 0. FAILURES OUT OF 3 TESTS
DOUBLE CANCELLATION IN S x T = 0. FAILURES OUT OF 3 TESTS

JOINT INDEPENDENCE

T x E INDEPENDENT OF S = 0. FAILURES OUT OF 108 TESTS
E x S INDEPENDENT OF T = 0. FAILURES OUT OF 108 TESTS
S x T INDEPENDENT OF E = 0. FAILURES OUT OF 108 TESTS

OPTIONS - PROTOTYPES

F1 GOTO NEXT PROTOYPE
F2 GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
F3 PRINT SUMMARY OF AXIOM VIOLATIONS
F4 PRINT COMPLETE AXIOM HISTORY
ESC MAIN MENU

Figure 8. Summary of Axiom Violations for the Four Stress Prototype Subjects

Return to the Program Setup screen and choose No. 6, Individual Axioms, to investigate subjects No. 8 and No. 12 individually. Choose these two subjects on the subject selection menu and proceed. Figure 9 shows the axiom violations for subject No. 8 and indicates that the analysis for this subject falls right on the border of acceptability. Pressing F1 will display axiom information for subject No. 12, as in Figure 10. This analysis far exceeds the criterion of axiom violations, as indicated by a maximum of 40 violations. Now the question is how do we develop a scale for these two subjects.

An additive solution definitely is not feasible to represent subject No. 12's data; therefore, this subject would either have to be dropped from the study or asked to provide additional information. Refer back to Section 2.4.3 for further information on handling this type of problem. However, even though there is considerable inconsistency in subject No. 8's ordering, an additive model will appropriately represent this data set. The only problem is that if No. 12 is dropped, this leaves No. 8 alone in the effort group. Typically, the goal is to try to have as many subjects in a group as possible, since averaging tends to eliminate any random error that may be present in the data. However, referring back to Figure 5 indicates that No. 8 had an equally high correlation with the stress group, which suggests the solution of changing his/her prototype membership to stress and repeating the analysis. Figure 11 indicates the procedure for doing this. Return to the Prototype Correlations screen and enter F1 to change the prototype. Do this by moving the cursor to the desired subject and entering the proper letter. To exclude a subject, enter an "L" as the prototype group. Notice that subject No. 12 has been excluded by this technique. Upon pressing F1 again, the program will recalculate the Kendall's Coefficient and the correlations, as indicated by Figure 12. As is evident in this figure, the Kendall's increased to .758 with the deletion of No. 12 (the group Kendall's does not change by changing prototype group as for subject No. 8).

After changing the groupings, there are two other aspects which need to be checked. First of all, the Kendall's Coefficient for each of the prototype groups (in this case only time and stress) should be calculated. The procedure for doing this is as just described; delete all subjects which are not in a particular prototype group. The results for the time group are shown in Figure 13. This figure indicates a Kendall's of .8835, well above the criterion. Similarly for the stress group, Figure 14 shows that the Kendall's is .9119. Secondly, since subject No. 8 has been added to the stress group, the axiom violations for this new group should be checked. This is done by selecting "4" on the Program Setup screen. As presented in Figure 15, adding this extra subject only slightly changes the axiom violations, and the solution for this group is certainly still acceptable.

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                INDIVIDUAL ANALYSIS
                    SUBJECT #  8

          INDEPENDENCE
T INDEPENDENT OF E AND S =  16. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =   0. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =  16. FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   1. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN E x S =   2. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   1  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =  14. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =  12. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =  14. FAILURES OUT OF 108  TESTS


                  OPTIONS - INDIVIDUAL
           F1   GO TO NEXT INDIVIDUAL
           F2   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
           F3   PRINT SUMMARY OF AXIOM VIOLATIONS
           F4   PRINT COMPLETE AXIOM HISTORY
          ESC   MAIN MENU
```

Figure 9. Summary of Axiom Violations for Su ject No. 8--Individually

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                INDIVIDUAL ANALYSIS
                   SUBJECT # 12

          INDEPENDENCE
T INDEPENDENT OF E AND S =  34. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =  34. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =  36. FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF   1  TESTS
DOUBLE CANCELLATION IN E x S =   2. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   0  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =  36. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =  18. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =  40. FAILURES OUT OF 108  TESTS


                  OPTIONS - INDIVIDUAL
        F1   GO TO NEXT INDIVIDUAL
        F2   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
        F3   PRINT SUMMARY OF AXIOM VIOLATIONS
        F4   PRINT COMPLETE AXIOM HISTORY
        ESC  MAIN MENU
```

Figure 10. Summary of Axiom Violations for Subject No. 12--Individually

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .7380

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

SUBJECTS 12 PRESS F1 TO QUIT

SUB. #	TES	TSE	ETS	EST	SET	STE	SUGGESTED PROTOTYPE
1	.80	.80	.66	.61	.59	.64	T
2	.80	.79	.74	.71	.68	.70	T
3	.30	.43	.43	.60	1.00	.96	S
4	1.00	.95	.60	.42	.27	.41	T
5	.43	.60	.30	.43	.96	1.00	S
6	.79	.84	.60	.58	.74	.80	T
7	.42	.59	.33	.46	.97	1.00	S
8	.64	.64	.75	.78	.78	.74	S
9	.99	.97	.56	.40	.35	.49	T
10	.97	.94	.68	.56	.48	.58	T
11	.43	.60	.30	.43	.96	1.00	S
12	.45	.42	.70	.75	.67	.58	L

Figure 11. Prototype Analysis Screen Indicating Procedure for Excluding Subject No. 12

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .7580

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

	SUBJECTS 12						SUGGESTED
SUB. #	TES	TSE	ETS	EST	SET	STE	PROTOTYPE
1	.80	.80	.66	.61	.59	.64	T
2	.80	.79	.74	.71	.68	.70	T
3	.30	.43	.43	.60	1.00	.96	S
4	1.00	.95	.60	.42	.27	.41	T
5	.43	.60	.30	.43	.96	1.00	S
6	.79	.84	.60	.58	.74	.80	T
7	.42	.59	.33	.46	.97	1.00	S
8	.64	.64	.75	.78	.78	.74	S
9	.99	.97	.56	.40	.35	.49	T
10	.97	.94	.68	.56	.48	.58	T
11	.43	.60	.30	.43	.96	1.00	S
12	.45	.42	.70	.75	.67	.58	L

F1 CHANGE PROTOTYPE
F2 PRINT
F3 RETURN TO PROGRAM SETUP
F4 GO TO NEXT OPTION CHOOSEN IN PROGRAM SETUP
ESC MAIN MENU

Figure 12. Prototype Analysis Screen Indicating Recalculation of the Kendall's Coefficient Excluding Subject No. 12

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .8835

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

SUBJECTS 12

SUB. #	TES	TSE	ETS	EST	SET	STE	SUGGESTED PROTOTYPE
1	.80	.80	.66	.61	.59	.64	T
2	.80	.79	.74	.71	.68	.70	T
3	.30	.43	.43	.60	1.00	.96	L
4	1.00	.95	.60	.42	.27	.41	T
5	.43	.60	.30	.43	.96	1.00	L
6	.79	.84	.60	.58	.74	.80	T
7	.42	.59	.33	.46	.97	1.00	L
8	.64	.64	.75	.78	.78	.74	L
9	.99	.97	.56	.40	.35	.49	T
10	.97	.94	.68	.56	.48	.58	T
11	.43	.60	.30	.43	.96	1.00	L
12	.45	.42	.70	.75	.67	.58	L

F1 CHANGE PROTOTYPE
F2 PRINT
F3 RETURN TO PROGRAM SETUP
F4 GO TO NEXT OPTION CHOOSEN IN PROGRAM SETUP
ESC MAIN MENU

Figure 13. Prototype Analysis Screen Indicating Recalculation of the Kendall's Coefficient Including Only the Time Prototype Subjects

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .9119

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

	SUBJECTS 12						SUGGESTED
SUB. #	TES	TSE	ETS	EST	SET	STE	PROTOTYPE
1	.80	.80	.66	.61	.59	.64	L
2	.80	.79	.74	.71	.68	.70	L
3	.30	.43	.43	.60	1.00	.96	S
4	1.00	.95	.60	.42	.27	.41	L
5	.43	.60	.30	.43	.96	1.00	S
6	.79	.84	.60	.58	.74	.80	L
7	.42	.59	.33	.46	.97	1.00	S
8	.64	.64	.75	.78	.78	.74	S
9	.99	.97	.56	.40	.35	.49	L
10	.97	.94	.68	.56	.48	.58	L
11	.43	.60	.30	.43	.96	1.00	S
12	.45	.42	.70	.75	.67	.58	L

F1 CHANGE PROTOTYPE
F2 PRINT
F3 RETURN TO PROGRAM SETUP
F4 GO TO NEXT OPTION CHOOSEN IN PROGRAM SETUP
ESC MAIN MENU

Figure 14. Prototype Analysis Screen Indicating Recalcualtion of the Kendall's Coefficient for the Stress Prototype Subjects and Subject No. 8

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                    PROTOTYPE ANALYSIS
                      STRESS PROTOTYPE

          INDEPENDENCE
T INDEPENDENT OF E AND S =   0. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =   0. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =   0. FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   1. FAILURES OUT OF   2  TESTS
DOUBLE CANCELLATION IN E x S =   0. FAILURES OUT OF   3  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   3  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =   8. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =   0. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =   0. FAILURES OUT OF 108  TESTS

                    OPTIONS - PROTOTYPES
              F1  GOTO NEXT PROTOYPE
              F2  GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
              F3  PRINT SUMMARY OF AXIOM VIOLATIONS
              F4  PRINT COMPLETE AXIOM HISTORY
              ESC MAIN MENU
```

Figure 15. Summary of Axiom Violations for the Modified Stress Prototype Group

Now that all of the potential problems have been addressed, scaling solutions can be generated for the prototype groups. Make sure to maintain the suggested prototypes as described above (subject No. 8 as a stress and subject No. 12 omitted). Then, select "5" on the Program Setup screen to obtain scaling solutions for the two prototype groups. The Scaling Information screen should appear as in Figure 16 for the time group and as in Figure 17 for the stress group. Selecting F2 will print the vital information for permanent record of the solutions. These values should be used (for the appropriate subjects) as the workload levels for the corresponding ratings from the Event Scoring phase. Event Scoring will be described in more detail in the next section. A similar procedure should be followed for every group of card sort data. While the exact approach will vary for each group, this general framework should prove successful.

******* SCALING INFORMATION *******

TIME SCALE

LAST 5 ITERATIONS

ITERATION	THETA	TAU
15	.01456	.94302
16	.01458	.94872
17	.01458	.94302
18	.01458	.94872
19	.01458	.93162

APPROXIMATE RELATIVE IMPORTANCE
OF EACH FACTOR

50.00 % FOR FACTOR T
28.07 % FOR FACTOR E
21.94 % FOR FACTOR S

THE SCALE VALUES FOR THE ITERATIONS BELOW
ARE PRINTED FROM ITERATION NO. 19

VARIABLE	ADDITIVE MODEL	ADDITIVE RESCALED
1 TIME 1	-.57	-8.23
2 TIME 2	-.01	15.86
3 TIME 3	.59	41.77
4 EFFORT 1	-.32	2.73
5 EFFORT 2	-.01	15.87
6 EFFORT 3	.33	30.80
7 STRESS 1	-.25	5.49
8 STRESS 2	.00	16.48
9 STRESS 3	.25	27.43

OPTIONS - PROTOTYPE

F1 PLOT OF RESCALED VS. RAW DATA
F2 PRINT SCALING INFORMATION
F3 PRINT ALL ITERATIONS
F4 VIEW SCALING SOLUTION
F5 GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
F6 GOTO NEXT PROTOTYPE
ESC MAIN MENU

Figure 16. Scaling Information for the Six Time Prototype Subjects

```
                 *******  SCALING INFORMATION  *******
                             STRESS SCALE
  LAST 5 ITERATIONS                  THE SCALE VALUES FOR THE ITERATIONS BELOW
ITERATION  THETA   TAU                ARE PRINTED FROM ITERATION NO.    14
  10       .01132  .94872
  11       .01134  .94302                       ADDITIVE     ADDITIVE
  12       .01134  .94872             VARIABLE   MODEL       RESCALED
  13       .01133  .94302             1   TIME 1    -.20         6.94
  14       .01133  .94872             2   TIME 2    -.02        15.48
                                      3   TIME 3     .22        26.58
                                      4 EFFORT 1    -.18         7.74
APPROXIMATE RELATIVE IMPORTANCE       5 EFFORT 2    -.02        15.28
       OF EACH FACTOR                 6 EFFORT 3     .20        25.98
                                      7 STRESS 1    -.65       -14.68
  15.64  % FOR FACTOR T               8 STRESS 2     .00        16.24
  18.23  % FOR FACTOR E               9 STRESS 3     .65        47.44
  62.12  % FOR FACTOR S
                                   OPTIONS - PROTOTYPE
                      F1   PLOT OF RESCALED VS. RAW DATA
                      F2   PRINT SCALING INFORMATION
                      F3   PRINT ALL ITERATIONS
                      F4   VIEW SCALING SOLUTION
                      F5   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
                      F6   GOTO NEXT PROTOTYPE
                      ESC  MAIN MENU
```

Figure 17. Scaling Information for the Five Stress Prototype Subjects

Section 3
EVENT SCORING

3.1. INTRODUCTION

The second phase of SWAT, Event Scoring, is the experiment or situation that the investigator is interested in evaluating for workload. In this situation, a subject performs a task and the investigator obtains information about the amount of workload associated with task accomplishment. When SWAT is used, this type of information is obtained by applying the same set of descriptors that the subject has become familiar with during the Scale Development phase. After completion of a task or task segment, the subject is asked to give a rating for Time Load, Mental Effort Load, and Psychological Stress Load. The subject responds by giving either a 1, 2, or 3 for each of the three dimensions. The three levels are defined in the same way that they were for the composition of the cards used in Scale Development. The difference is that, in the Scale Development phase, the subject had to recall some event from his or her past experiences for which the particular combination of descriptors was representative. On the other hand, in the Event Scoring phase, the subject experiences the event (task) and must select the correct set of descriptors that describes the workload created by the event.

If a considerable amount of time has elapsed between the Scale Development (card sort) and the study, it is important to provide refresher instructions to reacquaint the subjects with the descriptors that are used to rate the events during the experiment. An example of refresher instructions is included in Appendix G. Also included in Appendix G is a "SWAT Dimensions" card. Copies of this card may be used by subjects as a quick reference card during event scoring. If experimental conditions do not permit a subject to refer to the card at the time of giving event ratings, the card can be used as a last minute refresher just prior to beginning the experiment.

When referring to the SWAT dimensions or obtaining ratings on the dimensions, the order of Time Load, Mental Effort Load, and Psychological Stress Load should always be maintained. This convention has been established in order to minimize the mental load associated with communicating about the dimensions. Such standardization allows the subjects to devote less attention to deciding which dimension to evaluate first, second, and third, thereby reducing the chances of the rating task being intrusive on primary task performance. Acton and Colle (1984) addressed the issue of order effects by investigating the relationship between dimension order and ratings. They found that the interaction between order of dimensions used to obtain ratings and the workload level presented to the subject was not statistically significant. This result can be interpreted to indicate that the order of presentation of dimensions does not affect the ratings assigned to task levels.

3.2. TASK DEFINITION

3.2.1. Laboratory Tests

Experiments in the laboratory are generally divided into a series of trials. It is then a simple matter to obtain ratings between trials. The exact method of obtaining the ratings may vary to fit the constraints of each experiment. Ratings may be spoken, written, or entered into the experimental apparatus. The cue for when the subject is to respond may be provided by the experimenter or automatically presented by the apparatus. Since laboratory tasks tend to be abstractions of "real-world" tasks, it is important to give subjects enough training in the range of tasks to be used in the experiment so that he or she has some idea of the meaning of the various levels of the task. Practice in giving ratings should also be given to eliminate interference or erratic ratings due to unfamiliarity with the experimental procedures.

3.2.2. Operational Evaluations

Operational evaluations generally do not have the convenient break characteristic of laboratory trials. Operational tasks are generally continuous; in many cases, the evaluation will last several hours. In these cases, component tasks or subtasks will most likely be of interest to the investigator. The investigator should identify the subtasks of interest through either a formal or informal task analysis. Tasks should be identified in a way that is meaningful to the operators to facilitate their ability in identifying the task that they are trying to evaluate. In other words, a pilot would have very little problem in relating to the kind of load he was under from the time he passed the middle marker to touchdown, while he would have a very difficult time relating what kind of load he was under for 3 minutes after being presented a tone. One weakness in selecting the tasks for operational meaningfulness is that each of the tasks may have a different duration. In the previous example, the landing task might have a duration of approximately 45 seconds while lift off to passing through 10,000 feet might be over 10 minutes. Some psychometricians would say that, because of the difference in duration, measures of these tasks are not comparable. While there is undoubtedly noise injected into these data due to this difference, such noise very likely is less severe than the noise injected by artificially segmenting "real-world" tasks. As in laboratory tasks, ratings can be obtained either verbally or through some automated data collection process. Likewise, the cue to prompt the subject to give a rating can be either an inquiry by the experimenter or a signal of some type that is under equipment control. In some cases, there are capabilities inherent in the operational equipment that can be adapted to this purpose.

A primary concern in operational tests is to obtain data without interfering with the operator's ability to perform the task. Most likely, if the entire complex task or job is considered, a very large number of task segments will exist. If the investigator requires a response for every identifiable task segment, then the objective of not interfering with the operator's task performance will probably not be attained. Therefore, it is important for the investigator to identify the crucial segments to provide the needed data. This may include some "standard" events as well as the events where task overload is suspected. It should be remembered that, at least at this time, the measurement of workload can be attained only in a relative sense. In other words, if we take measurements, we can say that Task A has a workload of 88, while Task B has a workload of 37, for instance. Consequently, we are able to state which task has the highest workload and with certain measurement tools (SWAT, for example) we can say how much more workload one task has than the other. This relative nature of the current measurement tools requires the use of a comparison or baseline task. For example, if the investigator is concerned that the use of a particular automatic landing system has an excessive workload associated with its use then data might also be collected employing a system that is already in service. It is assumed that the new system would have increased capabilities associated with it but the increased capability could also place increased demands upon the operator. Data that are descriptive of the current system would give system designers and evaluators the information necessary to trade off cost against increased capability. Clearly the investigator must be careful in these kinds of evaluations to equate training to the maximum degree possible. Another aspect of attempting to remain unintrusive pertains to when the ratings must be obtained relative to the segment of interest. What is generally known about the fallibility of the human memory would seem to suggest that ratings should be taken as close to the event being rated as possible. If, for example, the investigator desires to obtain data on the aircraft approach task, he might want to obtain ratings as the aircraft passes the outer marker, the middle marker, and at touchdown. However, operational constraints or safety of flight considerations might preclude obtaining ratings at each of these points and may dictate that this recommendation be altered. One way that might be considered in attempting to get data in these kinds of demanding situations is to give the operator the option of not responding. The operator could be instructed to omit the response any time that workload is so high that to attempt a response would interfere with task performance. Then, for every task segment that the operator failed to rate, the investigator would enter a 3,3,3 or 100 in the data matrix since the operator was at the defined maximum workload level.

An alternative approach to this problem would be to delay data collection until a convenient break. In the example, the three ratings for the landing might all be taken after the airplane has finished the landing roll and the pilot has taxied off the active runway. This approach suggests the question, how long can the investigator wait to obtain the ratings?

Several investigations have been directed toward answering this question. Eggemeier, Crabtree, and LaPointe (1983) found that SWAT ratings delayed 15 minutes after completing a short-term memory task did not differ significantly from immediate ratings. Eggemeier, Melville, and Crabtree (1984) investigated the effect of an intervening task on delayed SWAT ratings in a short-term memory task. Their results indicate that, while there was not a significant effect of a 14-minute delay or intervening task on SWAT ratings, there was a tendency for ratings to be affected by performing a very difficult intervening task. Notestine (1984) found that a delay of 30 minutes did not significantly affect SWAT ratings in a probability monitoring task. In these last two experiments, however, subjects tended to report lower SWAT ratings after a delay than immediately after task completion. In many situations, such as some flight tests, it may be necessary to obtain all of the data after the entire task or mission has been completed. In these debriefing sessions, questionnaires are frequently used to obtain operator opinion data. If the ratings MUST be obtained post hoc, the best prompts available should be used to help the operator recall what he/she was experiencing at the moment of interest (Arbak, Shew, and Simons, 1984). In some cases, the operators have been shown videotapes of the events and asked to give the ratings in this prompted post hoc fashion. This approach is probably the best of these prompts. If the investigator remembers to obtain the baseline tasks in the same manner, then whatever effect on the ratings that exist that might be attributable to the post hoc rating approach is a constant across all tasks and, thus, will not affect the relative standing. However, the investigator should remember how the data were obtained when interpreting the data.

3.3. DATA ANALYSIS

After the event scoring phase has been completed and all ratings have been obtained, rescaled values for each of the ratings must be assigned. These rescaled values were generated by the scale development phase as described earlier. When assigning rescaled values, it is imperative that the appropriate scale be used for each individual. If a single group solution was deemed appropriate, it may be used for all of the subjects.

However, if subgroup or individual solutions were necessary, the proper scaling solution must be used for each subject (e.g., if subject No. 1 is a time prototype subject and prototype group solutions are being used, then use the scaling solution for the time group in assigning subject No. 1's rescaled values).

As an example of this procedure, suppose that a given task contained ten events for which SWAT ratings were obtained. For simplicity, only a single subject will be described. For the first event, the subject was under moderate time pressure, had a moderate amount of memory demands and

calculations, and very little psychological stress load. This subject most likely would give the three dimensions ratings of 2, 2, and 1, respectively. During scale development, this combination of levels could have been assigned a scale value of, for example, 14.1 based on the position of the 2-2-1 combination relative to all of the other 26 combinations represented in the card deck. Now that this combination has been selected as being descriptive of the experienced event, the 14.1 is entered into the investigator's data matrix as the workload scale value associated with performance of that event, by that subject, under that set of conditions. Table 3 graphically demonstrates how rescaled values are assigned to ratings. Based on this example, event number one for this subject would be assigned a value of 14.1. Similarly, event number two would be assigned a rescaled value of 57.9, event number three would receive a rescaled value of 0, and so on until all ratings have been assigned rescaled values. From this point on, data analysis is situation specific and will conform to the type of experiment being conducted.

TABLE 3. EXAMPLE OF HOW TO ASSIGN A RESCALED VALUE TO A SUBJECT'S GIVEN RATING

Event		SWAT Scale	
Event	Rating	Card Combination	Rescaled Value
1	2-2-1	111	0.0
2	2-1-3	112	24.4
3	1-1-1	113	51.4
4	3-1-2	121	7.6
5	1-3-3	122	32.0
6	2-3-1	123	59.0
7	2-1-2	131	27.7
8	1-2-2	132	52.1
9	3-3-3	133	79.1
10	3-2-1	211	6.5
		212	30.9
		213	57.9
		221	14.1
		222	38.5
		223	65.5
		231	34.2
		233	85.6
		311	20.9
		312	45.2
		313	72.3
		321	28.5
		322	52.9
		323	79.9
		331	48.6
		332	73.0
		333	100.0

Section 4
SUMMARY

If reviewed in its entirety, this user's guide should provide: (1) an overview of the tenet of mental workload; (2) a step-by-step description of how to implement SWAT as a measurement tool; and (3) step-by-step instructions for running the analysis software. The additional information on the conjoint measurement model including axiom testing and scaling was introduced to provide the technical background and reasoning underlying the application of conjoint measurement to SWAT.

Appendix A

CARD DECK FACSIMILE

A

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate Performance.

B

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

C

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

D

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

E

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

F

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

G

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

H

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

I

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

J

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

K

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

L

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

M

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

N

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

O

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

P

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

Q

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

R

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

S

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

T

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

U

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

V

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

W

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

X

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Y

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Z

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

ZZ

Occasionally have spare time. Interruptions or overlap among activities occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Appendix B
OUTLINE FOR VERBAL CARD SORT INSTRUCTIONS: COLLEGE STUDENTS

I. Introduction

1. Describe Concept of Workload

When we speak of "mental workload," we are clearly referring in some sense to mental effort. Our ideas about mental workload are affected by our experiences with physical workload. We can easily think of the effort one must expend to lift a heavy object, dig a ditch, or participate in our favorite sport. Physical work has been quantified in many ways including carbon dioxide production, heart rate, or amount of work performed in a unit of time. Mental workload, on the other hand, has proven to be more difficult to measure because it is something that occurs within the person and isn't directly observable.

We might think of mental workload as the amount of concentration required to write a paper, work simple addition problems, or solve complex algebra problems. There is probably unanimous agreement that the amount of work required to solve complex algebra problems would be greater than the amount of work required to solve simple addition problems. While some mental tasks are clearly "harder" than others, measures which quantify this phenomenon have been difficult to develop and validate.

2. Describe Purpose of This Experiment

The experiment you are participating in is concerned with mental, not physical workload, and we will deal with methods of measuring the amount of workload experienced while performing a computer-controlled laboratory task.

3. Performance as a Measure

There are several ways in which your mental workload could be measured, the first of which is to assess your performance on the task. For example, if the task is driving a car, the precision of following a desired track or reaction time to something suddenly entering your visual field, etc., could be an indicator of your workload. Frequently, however, there isn't any change in this type of observable performance, although two

people (or the same person under two different conditions) may experience differing degrees of effort expenditure to achieve this performance.

4. Discuss Subjective Measures

A. Unidimensional

Another way to attempt to measure your mental workload would be to simply ask you to rate your workload on a scale, say from one to ten, for whatever task you are performing. If daydreaming was labeled as "1" and intense concentration as "10," you could probably give a rating corresponding to the workload you were experiencing in performing a task. However, this approach does not give us much information about WHY you gave a particular rating, and we would not be sure that each person intended to describe the same level of workload even though the numbers used were the same.

B. Multidimensional

Another approach is to break up mental workload into several dimensions, or factors, which are generally considered to comprise workload. In this approach, you would not actually be giving ratings on workload, but you would rate the amount or degree of each factor that exists in a given task situation. An application of this approach has been developed and is called the Subjective Workload Assessment Technique (SWAT). This approach has an added feature of obtaining information from the people using the scale about how the identified factors go together to create their perception of mental workload.

II. Describe SWAT

1. Basic Concept

The technique describes subjective workload as being composed primarily of three dimensions: Time Load, Mental Effort Load, and Psychological Stress Load.

2. Brief Description of the Dimensions

It is important that you understand the meaning associated with the three dimensions and how they relate to the definition of workload. Let's go into a little more detail about the dimensions.

A. Time Load

(1) Description

Time Load refers to the amount of time pressure experienced in performing your task. This includes the fraction of total available time that you are busy and the degree to which different aspects of the task overlap or interfere with one another. Under high amounts of time load, you are unable to complete the task due to a shortage of time or interference created by the overlap activities.

(2) Example in Everyday Life

In a classroom test situation, there could be a high degree of Time Load caused by having a large number of problems (e.g., 100 versus 10) to solve and in a very limited amount of time (e.g., 30 minutes). Notice that we are not considering anything about how much effort is involved in solving the problems or the stress level involved in this situation.

B. Mental Effort Load

(1) Description

Mental Effort Load refers to the amount of attention and/or concentration required to perform a task. Tasks that require Mental Effort Load include storing and recalling things from memory, decision making, calculations, and problem solving. High levels of Mental Effort Load are required in situations that demand total concentration, whereas lower levels of Mental Effort Load are required when your mind wanders or your attention is distributed over more than one "easy" task component.

(2) Example in Everyday Life

Mental Effort Load could involve memorizing items, performing calculations on numbers, concentrating on listening to a speaker for important points, or making very difficult decisions. In the test situation, the problems to solve could be very difficult, requiring you to remember a formula, conversions, and complicated solution procedures. Or, they could be very easy with the solution to the problem being immediately obvious. The difficulty of the problems is not necessarily related to the amount of time provided to complete the test.

C. Psychological Stress Load

(1) Description

Psychological Stress Load refers to the presence of confusion, frustration, and/or anxiety which hinders completion of your task.

(2) Example in Everyday Life

Psychological Stress Load includes such things as pressure to excel, anxiety over physical dangers, tension, fatigue, general state of health or feelings, and comfort factors such as temperature or noise. In the test situation, if your course grade was to be determined by your performance on a certain test, there would probably be quite a high level of stress. However, in a situation where your grade was fairly well determined, the stress level would undoubtedly be less, regardless of the time pressure or the amount of concentration required. Also, if construction was going on near the test room, noise and distractions could affect your ability to concentrate and therefore impose psychological stress. In a driving situation, stress could be produced by obscure road signs, heavy traffic, or inclement weather, which could cause you to become lost, frustrated, or concerned for your safety.

3. Description of the Levels Within the Dimensions

Figure 18 presents the three dimensions which have just been described as the main contributors to workload and the levels of each dimension. Notice that there are three levels of each dimension which can be used to give a rating. One is associated with the lowest degree of each of the dimensions, three is associated with the highest degree and two is a middle degree. Verbal descriptors are provided to define how you should evaluate the levels of each of the dimensions.

4. Identification of Other Dimensions

While it can be seen that these three dimensions contribute to mental workload, you may be able to think of other dimensions that may have an effect on the workload involved in performing a task. While this may be true, we believe that these three dimensions can be used to cover most of what most people are referring to when they speak of workload. A more precise breakdown into more fundamental components would complicate the rating process and possibly interfere with the rater's ability to perform the assigned task thus contributing to workload.

5. Examples of the Interactions Between the Three Dimensions

Figure 19 provides an example of two situations and how workload may be affected. In the left scene, we have a pilot approaching an airfield on a clear, sunny day. Let us presume that he/she has made this landing many times before. In this situation for this pilot, the task might get a rating of 1 for Time Load, 1 for Mental Effort Load, and 1 for Psychological Stress Load. In contrast, on the right side, the situation has changed and the approach is being attempted in adverse weather with visibility obscured. Now the pilot might give a rating of 2 for Time Load since he is not able to see what is happening and is not able to perform the necessary actions as early as in the previous situation. He may give a rating of 1 again for Mental Effort Load, since the decisions and amount of concentration are very similar to the first situation. However, the Psychological Stress Load rating may increase to a 3 since there could be increased anxiety about the landing due to the obscured visibility. Note that these ratings are intended to illustrate what a pilot might say and that for any given task any other combination of ratings is possible depending upon the precise task conditions and pilot-related factors such as amount of training.

SWAT

I. TIME LOAD

1. OFTEN HAVE SPARE TIME. INTERRUPTIONS OR OVERLAP AMONG ACTIVITIES OCCUR INFREQUENTLY OR NOT AT ALL.
2. OCCASIONALLY HAVE SPARE TIME. INTERRUPTIONS OR OVERLAP AMONG ACTIVITIES OCCUR FREQUENTLY.
3. ALMOST NEVER HAVE SPARE TIME. INTERRUPTIONS OR OVERLAP AMONG ACTIVITIES ARE VERY FREQUENT OR OCCUR ALL THE TIME.

II. MENTAL EFFORT LOAD

1. VERY LITTLE CONSCIOUS MENTAL EFFORT OR CONCENTRATION REQUIRED. ACTIVITY IS ALMOST AUTOMATIC, REQUIRING LITTLE OR NO ATTENTION.
2. MODERATE CONSCIOUS MENTAL EFFORT OR CONCENTRATION REQUIRED. COMPLEXITY OF ACTIVITY IS MODERATLY HIGH DUE TO UNCERTAINTY, UNPREDICTABILITY, OR UNFAMILIARITY. CONSIDERABLE ATTENTION IS REQUIRED.
3. EXTENSIVE MENTAL EFFORT AND CONCENTRATION ARE NECESSARY. VERY COMPLEX ACTIVITY REQUIRING TOTAL ATTENTION.

III. PSYCHOLOGICAL STRESS LOAD

1. LITTLE CONFUSION, RISK, FRUSTRATION, OR ANXIETY AND CAN EASILY BE ACCOMMODATED.
2. MODERATE STRESS DUE TO CONFUSION, FRUSTRATION, OR ANXIETY NOTICEABLY ADDS TO WORKLOAD. SIGNIFICANT COMPENSATION IS REQUIRED TO MAINTAIN ADEQUATE PERFORMANCE.
3. HIGH TO VERY INTENSE STRESS TO DUE TO CONFUSION, FRUSTRATION, OR ANXIETY. HIGH TO EXTREME DETERMINATION AND SELF-CONTROL REQUIRED.

Figure 18. SWAT Rating Scales

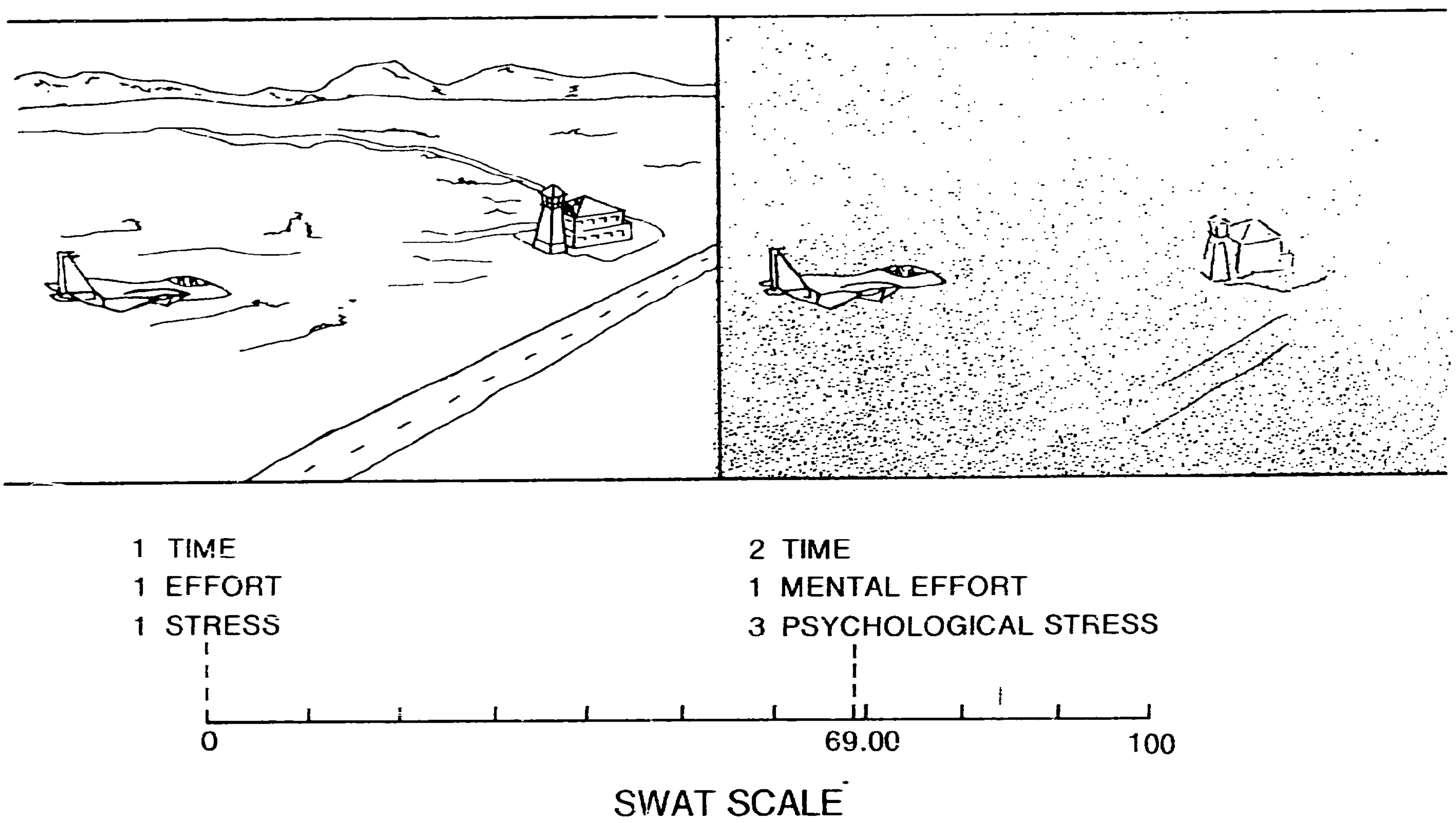

Figure 19. Event Scoring

6. Describe How Dimensions are Combined Into One Dimensional Scale

Another interesting aspect of SWAT is that the procedure provides a mathematical technique for combining your ratings on the three dimensions into a single workload scale, similar to a one to ten scale described earlier. As can be seen in these two situations, the 1-1-1 combination is translated into, naturally, the lowest possible scale value of 0. The 2-1-3 combination in the right scene is translated to a 69, for example.

Later you will perform a certain task, and give a rating for each of the three dimensions at the end of each task trial. These three numbers will then be combined as in this illustration into a value from 0 to 100. This will then serve as your workload score for that particular trial.

III. How Does This Relate to What We're Doing Today

For now, we need to find out information from you so that we can develop the mathematical model which will be used to combine the ratings you will give later. The information we are looking for concerns the importance you place on the three dimensions. You will tell us this by taking this deck of cards, each of which has a combination of the three dimensions on it, and order them from that combination which represents the lowest workload to the combination which represents the highest workload. Do this by imagining a situation that you have experienced which could be described by the combination of the dimensions on a particular card and making a relative judgment about the workload associated with accomplishing this task and rank the cards accordingly. As you can see, with three levels of each dimension there is a total of three times three times three, or 27 possible combinations which could be given. In this way, 27 cards are created which you must rank-order.

Notice that in doing this ranking procedure, the difficult part involves the trade-off decisions which have to be made. Suppose, for example, that you are comparing card 1-2-1 for time, effort, and stress, respectively, with card 1-1-2. Now in one situation the effort is higher, while in the other situation stress is higher. You must decide which situation you would choose as lower in workload. To do this, you need to decide which dimension, effort or stress, has a greater impact on the overall workload to you. Similar situations will arise with different combinations of all three dimensions. There is no right or wrong answer to these decisions, since each person feels differently about the importance of the dimensions. Some people feel that time has the greatest impact, others feel it is effort, and still others feel that stress is the most important.

Because we are trying to determine what constitutes mental workload for you, we would rather not supply you with examples of workload that are represented by each combination. We ask that you supply your own examples. This is necessary because a situation that is very demanding for one person might be very easy for another. Likewise, since we all have different backgrounds, we may or may not be able to relate to a specific example. For example, if I give examples that are related to flying an airplane, that could be very meaningful to a skilled and experienced pilot while most of you in this room might not have any such experience. Therefore, your impressions of the workload involved in such a task would be dependent upon impressions provided by other people.

To avoid this, you are asked to read the descriptors from a card and try to think of something you have experienced that this set of descriptors would have accurately described. Then take another card and repeat the procedure. By comparing the events which you have recalled determine which of the events had the highest workload for you. Repeat this process until you arrive at an order for the 27 cards that begins with the description of the lowest workload event and ends with the description of the highest workload event.

As you try to imagine situations which could be described by the combinations on the cards, there may be combinations for which you have a hard time imagining a situation that fits. In these cases, we could provide you with an appropriate situation, but it is more beneficial if you assume that such a situation or event does exist and try to determine where it would rank in relation to the other situations.

Pay attention to the verbal descriptors on the cards to make your judgments, as it is important that you become comfortable with the levels of the dimensions. This will help you later when you make your ratings.

Now read the written instructions which will describe in more detail the dimensions and the sorting task which you will do.

Appendix C

OUTLINE FOR VERBAL CARD SORT INSTRUCTIONS: PILOT POPULATION

I. Introduction

1. Describe Purpose of This Experiment

The experiment you are about to participate in is to evaluate the new XYZ all weather terrain following radar. It is very important for us to get the best data that we possibly can if we are to get this system delivered in a configuration that will require few modifications once it becomes operational. Considerable time and resources have already been expended to get these systems to this point. In this test, all of you will fly simulations over three different courses, using both candidate XYZ systems as well as one sortie that you will fly manually. These flights will be flown under simulated night as well as day conditions, and under three different kinds of simulated weather.

Performance will be recorded in terms of the precision with which you hold the assigned altitudes and course track, control actions, ordnance expended, and target scores. In addition to the performance data, we will be collecting data to determine the amount of workload associated with mission accomplishment.

2. Describe Concept of Workload

You are probably quite familiar with the concept of mental workload. This is a concept that has become increasingly important in modern high technology aircraft. When we speak of mental workload, we are referring to some sense of mental effort. The basic idea is that we have a finite capacity for performing mental work; and if we exceed this capacity, then we will begin to make a large number of errors or experience total performance breakdown. We can think of situations where very little effort is required thus leaving us with considerable spare capacity for work. Likewise, we can think of situations that require substantial effort leaving us with little or no spare capacity. Examples of performing the same task with varying amounts of workload include employing different systems to perform the same task, performing the same task with different levels of experience or skill, or performing the same task under different levels of factors related to the task such as fatigue, weather conditions, system malfunctions etc., that could affect the amount of a person's capacity that is required to maintain acceptable task performance.

In this study, we are going to measure workload through the use of a scaling approach called the Subjective Workload Assessment Technique or SWAT. This is a technique that has been developed and extensively tested at the Armstrong Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base and has been successfully used in a number of simulation tests, flight tests, and OT&Es. This technique is different from most scaling procedures in that there are two parts to it. The first part is called Scale Development and is what we are going to do today, and the second part is called Event Scoring and is the part where you will give workload ratings in the simulator. One of the primary objectives of this technique is to create as little interference as possible during task performance while getting the best quality data possible.

Before I start a more detailed explanation of the procedure, I would like for you to quickly read over the written instructions that you have been provided. Don't labor over these instructions since I am going to repeat much of it anyway. I want you to read the instructions to be sure I don't forget something important, and by getting a preview, it will help you follow what I am trying to tell you later and hopefully provide you with a better understanding of the procedure.

(PAUSE LONG ENOUGH FOR EVERYONE TO READ THE INSTRUCTIONS)

For the purposes of SWAT, workload has been defined as being composed primarily of three things: Time Load, Mental Effort Load, and Psychological Stress Load. Each of these three factors or dimensions has had three levels defined. Therefore, for the purposes of SWAT, workload can be represented by the cube shown in Figure 20. You can see that there are three levels of each of the three dimensions. All possible combinations of these dimensions comprise the 27 cells of this larger cube that we are calling workload. Your task today is, through a card sort procedure, to help us determine how these dimensions combine to create your conception of workload. The deck of cards in front of you has a card for each of the cells in this cube. Each card has three descriptors written on it--one for time load, one for mental effort load, and one for psychological stress load. By arranging this deck in an order that represents which combination you think describes the lowest workload condition to the combination that you think represents the highest workload condition and the 25 steps in between, you are helping us create a scale that will reflect the way you people (in this test) think these dimensions combine to create the impression of workload. This is not the same for everyone. Some people think that time is the only element that has any importance for determining

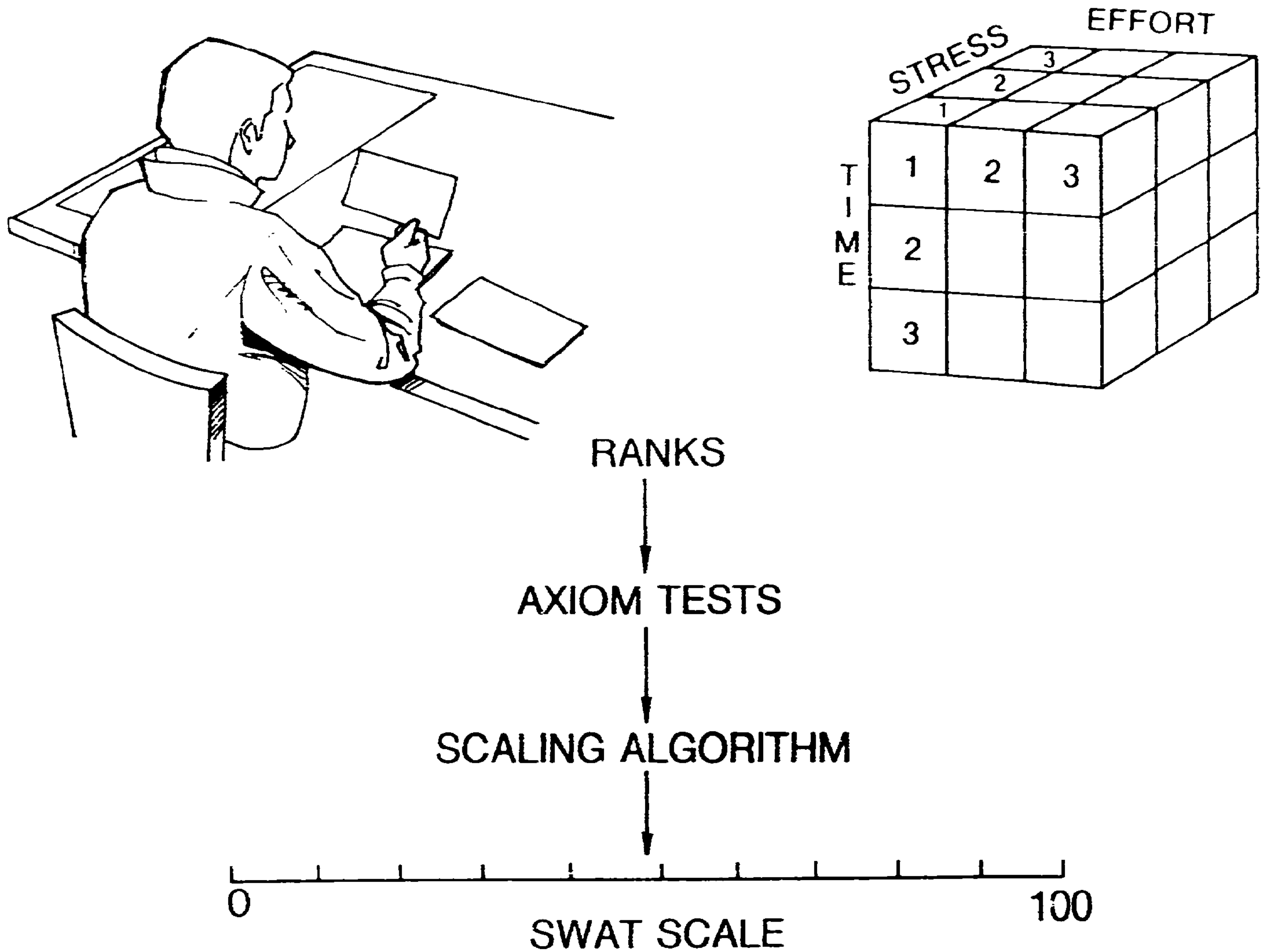

Figure 20. Scale Development

workload, while other people will say that the only thing that is important to them is to be able to manage the psychological stress, and others think that difficulty of the required tasks is most important.

3. Definition of Dimensions

As a start, you need to understand a little better what we mean by these three dimensions.

A. Time Load

(1) Description

Time Load is the amount of time pressure experienced in performing your task. This includes the fraction of total available time that you are busy and the degree to which different aspects of the task overlap or interfere with one another. Under high Time Load, you are unable to complete the task due to a shortage of time or interference created by overlap of activities.

(2) Example

For example, in an emergency situation, especially in a situation with multiple emergencies, the required actions may be relatively simple and well practiced. The only real problem may be that things happen so fast that you just cannot get everything accomplished before things go from bad to worse.

B. Mental Effort Load

(1) Description

Mental Effort Load is the amount of attention and/or concentration required to perform a task. Things that are considered mental effort include recalling things from long-term memory, decision making, performing calculations, storing and retrieving things from short-term memory, and problem solving. High levels of Mental Effort Load are required in a situation which demands total concentration to perform, while during lower levels of mental effort, your mind may wander or attention may easily be shared with several relatively easy tasks.

(2) Example

Mental Effort Load could involve such things as remembering a radio frequency that must be dialed in after passing some navigation point or having to make a decision regarding which of several potential targets should be attacked and what direction to approach a target from on each pass.

Another example might be the memory load associated with remembering a complex procedure needed to activate a particular piece of equipment. This situation might be intensified if employment of this piece of equipment is a rare event and, therefore, not as thoroughly learned as a more frequently occurring event.

C. Psychological Stress Load

(1) Description

Psychological Stress Load refers to the presence of confusion, frustration, and/or anxiety which hinders completion of your task. Psychological stress refers to the feelings of apprehension and tension one usually thinks of when you hear the term stress. In addition, other factors such as fatigue, motivation, and low levels of physical stressors may also contribute to the feeling of psychological stress load.

(2) Example

It is well known that physical stressors such as G forces, vibration, temperature and noise can, when existing in sufficient magnitude, interfere with task performance. At low levels, these stressors may not produce interference but will produce enough of an annoyance that after a certain amount of time some of the person's capacity may be expended just to keep the irritation pushed into the background. This level of capacity expenditure we would attribute to the psychological stress dimension of workload.

4. Description of Levels Within the Dimensions

Now that we have some idea what we mean by the three dimensions, notice in our cube that each dimension has three levels. Level one is associated with the lowest degree of each dimension, level three is associated with the highest degree of each dimension, and level two is associated with a moderate degree of load for each dimension. Verbal descriptors have been written to precisely define each of the levels for each of the dimensions. The numbers for these three dimensions will be used by you later when you are doing the event scoring. The descriptors have already been introduced to you when you read the written instructions. Now as you arrange the card deck in order from the lowest to the highest workload situation you will read each of these descriptors many times. This will help you to become familiar with the meaning associated with each level of each dimension. You are asked to try to think of the wording of the descriptors as you do your ordering rather than trying to use the numbers associated with the levels or in some other way attempt to mechanize the ordering. The ordering information is very important in helping to define the scale but equally important to us is the training value associated with carefully considering the relationships of the meaning of the levels for each of the three dimensions.

Several points need to be made at this stage:

- Remember that there is not a correct answer. You are making judgments about conditions in terms of the degree of workload associated with an event like the one under consideration. This is a communication process that we use for you to try to explain to us the way you view workload in terms that allow us to put numbers on your judgments. In this process, it may become necessary for us to later come back to you to confirm our opinion of what you are trying to tell us or resolve some area that is not clear to us. If this does become necessary, do not change your judgments because you think you got the "wrong answer." There is not a "right" and "wrong" answer. Just try to be consistent in giving your judgment about events.

- Because people do differ, it is best that you not "compare notes" with anyone. Do not discuss things in an attempt to come up with a consensus.

- As you do the card sort, try to think of an experience that you have had that each card (set of descriptors) would describe. Then put the cards in order by deciding which of the experiences had the higher workload. Remember, you provide the event--something with which you are familiar. If we specified events for you to rank, we very likely would describe

something with which you are not familiar. This would make your judgments of the relative workload less valid.

This process of recalling events helps to establish a scale that is representative of this group of subjects' opinion.

Some of the combinations may not remind you of a particular event. It may be very difficult to think of how you could have the highest level of one dimension while having the lowest level on the other two dimensions. It is true that in most cases the levels of the dimensions will go in the same direction. In other words, as Time Load becomes greater, then both Mental Effort Load and Psychological Stress Load are probably going to be increasing also and conversely. On the other hand, as we have discussed this with people, they have generally been able to think of events where the more extreme combinations have existed. We would suggest, therefore, that these combinations do exist but are not the most frequently occurring type of events. If you simply cannot think of an event that a particular combination of descriptors would be appropriate for, then you should think, "if an event did exist, where would it fall in this order." Also remember that we are asking you to judge "how much" work is associated--not which would you like to have. It might be clear that one task has lower workload associated with it than another. In fact, the first task could be so low in demand that in your judgment it would be boring. Someone who has a low tolerance for boredom might be tempted to think, "I know this task is low in workload but I really do not like to be bored so I will rank it higher than this other task." Remember it is not your preference we are asking for but the amount of workload you think exists in a situation.

You may use whatever strategy that seems best to you to do the card sort. The procedure of dividing the 27 card deck into three smaller decks of low, medium, and high and ordering these smaller decks and then recombining into the entire 27 is a strategy that has proven useful to many people. However, this strategy is not mandatory. This is not an easy task. It will probably take you from 30 minutes to an hour to finish and some of the discriminations are going to be difficult. Please concentrate and give us the best sort possible. Even though this is a laborious process, we think that it pays off in the long run. When you get to the simulator, the rating task will be easier because of the effort you are putting in now.

Are there any questions? Then, go ahead and start. If you have any questions at any time, please ask them. The thing that is bothering you may also be bothering someone else. If you are not certain about how to do the ordering, it will affect your data so please ask if you have any questions.

Appendix D
WRITTEN CARD SORT INSTRUCTIONS

SWAT CARD SORT INSTRUCTIONS FOR SUBJECTS

During the course of this experiment, you will be asked to quantify the mental workload required to complete the tasks you will be performing. Mental workload refers to how hard you work to accomplish some task, group of tasks, or an entire job. The workload imposed on you at any one time consists of a combination of various dimensions which contribute to the subjective feeling of workload. The Subjective Workload Assessment Technique (SWAT) defines these dimensions as (1) Time Load, (2) Mental Effort Load, and (3) Psychological Stress Load.

For the purposes of SWAT, the three dimensions have been assigned three levels. The dimensions and their levels are described in the following paragraphs.

Time Load

Time Load refers to the amount of spare time that you have available (fraction of the total time that you are busy). When Time Load is low, sufficient time is available to complete all of your mental work with some time to spare. As Time Load increases, spare time drops out and some aspects of performance overlap and tasks interrupt one another. This overlap and interruption can come from performing more than one task or from different aspects of performing the same task. At higher levels of Time Load, several aspects of performance often occur simultaneously, you are busy, and interruptions are very frequent.

Time Load may be rated on the three point scale below:

(1) Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

(2) Occasionally have spare time. Interruptions or overlap among activities occur frequently.

(3) Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Mental Effort Load

As described above, Time Load refers to the amount of time one has available to perform a task or tasks. In contrast, Mental Effort Load is an index of the amount of attention or mental effort required by a task regardless of the number of tasks to be performed or any time limitations. When Mental Effort Load is low, the concentration and attention required by a task is minimal and performance is nearly automatic. As the demand for mental effort increases due to task complexity or the amount of information which must be dealt with in order to perform adequately, the degree of concentration and attention required increases. High Mental Effort Load demands total attention or concentration due to task complexity or the amount of information that must be processed.

Mental Effort Load may be rated using the three point scale below:

(1) Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

(2) Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

(3) Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Psychological Stress Load

Psychological Stress Load refers to the contribution to total workload of any conditions that produce anxiety, frustration, or confusion while performing a task or tasks. At low levels of stress, one feels relatively relaxed. As stress increases, confusion, anxiety, or frustration increases and greater concentration and determination are required to maintain control of the situation.

Psychological Stress Load may be rated on the three point scale below:

(1) Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

(2) Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

(3) High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Each of the three dimensions just described contribute to workload during performance of a task or group of tasks. Note that, although all three factors may be correlated, they need not be. For example, one can have many tasks to perform in the time available (high Time Load) but the tasks may require little concentration (low Mental Effort Load). Likewise, one can be anxious and frustrated (high Psychological Stress Load) and have plenty of spare time between relatively simple tasks. Since the three dimensions contributing to workload are not necessarily correlated, please treat each dimension individually and give independent assessments of the Time Load, Mental Effort Load, and Psychological Stress Load that you experience in performing the following tasks.

One of the most important features of SWAT is its unique scoring system. SWAT uses a procedure to find separate scoring weights for each level of a dimension. Then, it determines a distinctive workload scale for each person or group. This scaling system greatly improves the precision of the workload ratings you will give later.

In order to develop your individual scale, we need information from you regarding the amount of workload you feel is imposed by various combinations of the dimensions described above. We get this information by having you rank order the workload associated with each of the combinations.

In order for you to rank order the workload for each of the combinations, you have been given a set of 27 cards with the combinations from each of the three dimensions. Each card contains a different combination of levels of Time Load, Mental Effort Load, and Psychological Stress Load. Your job is to sort the cards so that they are ranked according to the level of workload represented by each card.

In completing your card sorts, please consider the workload imposed on a person by the combination represented in each card. Arrange the cards from the lowest workload condition through the highest condition. You may use any strategy you choose to order the cards. One strategy that is useful is to arrange the cards into three preliminary stacks representing "high," "moderate," and "low" workload. Individual cards may be exchanged between stacks, if necessary, and then rank ordered within stacks. Stacks can then be recombined and checked to be sure that they represent your ranking of lowest to highest workload. However, the choice of strategy is up to you and you should choose the one that works best for you.

There is no "school solution" to this problem. There is no correct order. The correct order is what, in your judgment, best describes the progression of workload from lowest to highest for a general case rather than any specific event. That judgment differs for each of us. The letters you see on the back of the cards are to allow us to arrange the cards in a previously randomized sequence so that everyone gets the same order. If you examine your deck you will see the order on the back runs from A through Z and then ZZ.

Please remember:

(1) The card sort is being done so that a workload scale may be developed for you. This scale will have a distinct workload value for each possible combination of Time Load, Mental Effort Load, and Psychological Stress Load. The following example demonstrates the relationship between the card sort and the resulting workload scale:

<u>Time</u>	<u>Effort</u>	<u>Stress</u>	<u>Workload Scale</u>
1	1	1	0.0
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
3	3	3	100.0

(2) When performing the card sorts, use the descriptors printed on the cards. Please remember not to sort the cards based on one particular task (such as flying an airplane) or what you anticipate that you will be doing in this study. Sort the cards according to your general view of workload and how important you consider the dimensions of Time Load, Mental Effort Load, and Psychological Stress Load to be. Base these decisions on all types of experiences and task situations.

(3) During the actual experiment, you will accomplish the desired task. Then, you will provide a SWAT score based on your opinion of the mental workload required to perform the task. This SWAT score will consist of one number from each of the three dimensions. For example, a possible SWAT score is 1-2-2. This represents a 1 for Time Load, a 2 for Mental Effort Load, and a 2 for Psychological Stress Load.

(4) We are not asking for your preference concerning Time Load, Mental Effort Load, and Psychological Stress Load. Some people may prefer to be "busy" rather than "idle" in either Time Load, Mental Effort Load, or Psychological Stress Load dimension. We are not concerned with this preference. We need information on how the three dimensions and the three levels of each one will affect the level of workload as you see it. You may prefer a 2-2-2 situation instead of a 1-1-1 situation. However, you should still realize that the 1-1-1 situation imposes less workload on you and leaves a greater reserve capacity.

The card sort procedure will probably take 30 minutes to an hour. Please feel free to ask questions at any time. Thank you for your cooperation.

Appendix E

SWAT CARD SORT DATA SHEET

E/

M/

A/

N/

N												
B												
W												
F												
J												
C												
X												
S												
M												
U												
G												
Z												
V												

SWAT CARD SORT DATA SHEET (continued)

E												
M												
A												
N												
Q												
ZZ												
K												
E												
R												
H												
P												
D												
Y												
A												
O												
L												
T												
I												

Appendix F
INFORMATION FOR PROGRAM OPERATION

This appendix describes the requirements and methods for analyzing SWAT scale development data using the microcomputer-based software. The computer system requirements for running the program are:

- IBM PC or compatible system.
- 512K internal memory.
- Two floppy disks or one hard disk and one floppy disk.

The software requirements are:

SWATPRGM.BAT	This is the current version (3.1) of the SWAT program which will handle up to 30 subjects' data. To start program operation, type "SWATPRGM" and then return. This file is provided on the SWAT software disk.
MAIN.EXE	This is the executable program which is called by SWATPRGM. This file is provided on the SWAT software disk.
SWAT.DAT	This is a data file which provides parameter information to the program. This file is included on the SWAT software disk.
ANSI.SYS	This is a functional device driver that enables the program to use special character sequences. The ANSI driver MUST be added to the system's linked list of standard device drivers by inserting a DEVICE = ANSI.SYS in the CONFIG.SYS file in the root directory. Of course, the ANSI.SYS file must be in the root directory as well. Once you have done this, ANSI.SYS will be loaded into memory whenever you boot up your system. ANSI.SYS is part of the MS/DOS operation systems and therefore is purchased by each user.
Optional Math Coprocessor	The 8087 math coprocessor chip will speed up program execution but is not necessary to run the program on systems based on the 8086 or 8088 family of processors. The SWAT program will run on systems based on the 80286 family of processor chips as long as the system has a math

coprocessor included in the system configuration. SWAT will not run on a 80286-based computer system that does not have a math coprocessor.

During program operation, one file is written to the program disk. Because of this, the original disk is not write-protected. It is advisable to make backup disks immediately upon receipt of the program. We do ask, however, that you do not distribute the software to others. Since the SWAT program is government property, any individual can obtain a copy simply by writing the Harry G. Armstrong Aerospace Medical Research Laboratory. If copies are distributed without our knowledge, the recipient will not be advised of program updates or problems which may be discovered as the program is used.

The remainder of this appendix is organized for a complete description of each of the screens with which the user will interact. The program is interactive and the user menus and screens are designed to help guide you through the analysis process. The interface has been designed for ease of use and should facilitate data analysis. First, a sample screen will be presented for reference, and then the screen functions will be described. To the degree possible, the order of execution for a routine analysis will be followed. We will deviate from this practice only to completely describe all of the options presented on a given screen. Output information resulting from this analysis includes prototype correlations, axiom test results, and the conjoint scaling solution.

As an additional aid the (ESC) key has been programmed to be the user's "panic button." At anytime while you are working with the SWAT program that you become confused or lost in the analysis, you can press the (ESC) key and return to the Main Menu screen and all data will be retained by the program. In this way you can always get back to a place where you know your way out and avoid losing data.

EQUIPMENT SPECIFICATION

If this is the first time the program is used, an equipment specification screen as shown in Figure 21 will be displayed. The three options are:

1. Two floppy disk drives, A and B.
2. One floppy drive and one hard drive, A and E, respectively.
3. One floppy drive and one hard drive, A and C, respectively.

Make the selection which corresponds to the system being used. You will then be asked whether a printer is attached to the computer. Once this information has been entered, subsequent uses of the

```
*******  EQUIPMENT SPECIFICATION  *******

                OPTIONS:

1) 2 FLOPPY DISK DRIVES - A and B
     (source in drive A) and
     (data in drive B)

2) 1 FLOPPY DISK DRIVE   - A
   1 HARD DISK DRIVE     - E
     (source in drive E) and
     (data in drive A)

3) 1 FLOPPY DISK DRIVE   - A
   1 HARD DISK DRIVE     - C
     (source in drive C) and
     (data in drive A)

4)  EXIT WITH NO CHANGES

          MAKE A SELECTION:
```

Figure 21. Equipment Specification Screen

program will not invoke this menu. Changes in the equipment specifications can be made, however, using option 4 of the Main Menu.

MAIN MENU

As can be seen in Figure 22, this screen has three main functions. The first is to allow entering of date, study name, and additional comments for labeling a data set. When entering the date, include slashes as displayed to keep the month, date, and year separate. The study name, which can be used as a reference to the data set, can be up to 20 characters. The comment lines can be used for additional information such as subjects' names or other identifying comments. The second function of this screen is to specify the name of the file and number of subjects to be used to later identify the data set. When you assign the file name on this screen, it becomes the file the program uses when the data set is stored to disk. Consequently, the file name must be eight characters or less. The file name should be different from the study name or some variant of the study name to allow for multiple analyses you may do with the same data. As always, if you want to retrieve a data file for additional work, the file name you use must be typed in exactly as before. For example, you may want to access a previous data file to modify it by adding subjects. The number of

```
*******   COMMENTS AND MAIN MENU   *******

     TODAY'S DATE:
                   (mm/dd/yy)

     STUDY NAME:                           FILE NAME:
                 (20 CHARACTERS MAX)                   (8 CHARS. MAX)

     NUMBER OF SUBJECTS:

COMMENT:
COMMENT:
COMMENT:




** USE A SEPARATE DATA DISK FOR EACH STUDY **
```

Figure 22. Comments Section of Main Menu Screen

subjects will specify the allotment of space in the data entry screen. If you desire to work with an existing data file press RETURN for all entries on this screen EXCEPT filename. When you enter a file name of an existing file:

- The equipment specification will direct the system to the appropriate disk drive.

- You will be advised that the file already exists and asked to verify (with a "Yes" or "No") that you wish to use data from a previous file. This is a safeguard to prevent accidentally writing over an existing data file. The screen will change to an updated version which includes the information previously saved in that file, as depicted in Figure 23.

The third function of this screen is to select the program options. With the selection of either a new or an existing file, you will be asked to choose one of the following options to direct the program to specific operations.

```
******* COMMENTS AND MAIN MENU *******

TODAY'S DATE: 03/03/87
              (mm/dd/yy)

STUDY NAME:UG DATA SET              FILE NAME:SAMPLE
           (20 CHARACTERS MAX)                (8 CHARS. MAX)

NUMBER OF SUBJECTS:15

COMMENT:This is a sample data set to be used with the Users Guide
COMMENT:to demonstrate the structure of the program.
COMMENT:Typically these comment lines are used for study information

                            MAIN
                            MENU

      F1  EDIT COMMENTS                F4  EQUIPMENT SPECIFICATION
      F2  DATA ENTRY                   F5  END THE PROGRAM
      F3  PROGRAM SETUP

                       MAKE A SELECTION:
```

Figure 23. Main Menu Screen with Sample Comments

OPTIONS

F1, Edit Comments	This option allows you to change any or all of the information on the first screen, including the date, study name, file name, number of subjects, or additional comments. This can be used to retrieve a different already existing data set, for example, by changing the file name. Doing this will cause the program to search the disk for the newly entered file name. Once the program finds the file, a "yes" response to the prompt that asks whether to use the existing file will cause all of the stored information from that file to be loaded into the program and displayed in place of the current file. When you have the file you want in memory, you may use the Edit Comments option to change any piece of information on this screen. This is useful if you are doing additional analyses to a file previously analyzed, or if you are adding subjects. If you wish to add subjects simply edit the Number of Subjects specified on this screen. This more recent version of the data set can be saved under a different name (if both an old and new version are desired) through the Data Entry screen with the "Save Data" option. If you change any character on a line you must reenter the rest of the line, as the program will not save the characters to the right of the cursor when you press RETURN.
F2, Data Entry	This option takes you to the Data Entry screen to enter, add, subtract, or otherwise modify data.
F3, Program Setup	Selection of this option takes you to the Program Setup screen where you will select options that permit you to specify the analysis you want performed. See the description of the Program Setup screen for more details. This option allows you to bypass the Data Entry screen if data have already been entered.
F4, Equipment Specification	If you change systems, this option can be used to change the specifications to correspond to the present system.
F5, End the Program	After all analyses have been completed, this option allows you to exit to the system.

As can be seen, this screen allows you access to all parts of the program. Consequently, you will be returning here frequently, and this is where an (ESC) keypress will take you.

DATA ENTRY SCREEN

Upon selection of F2, Data Entry, on the Main Menu screen, a formatted screen as depicted in Figure 24 will be displayed that allows you to enter or modify data. The number of subjects that was previously entered on the Main Menu will be displayed at the top of the screen, and sufficient space for data for this number of subjects will be created. While decimals are not typically used, space is included for special applications which require their use. Normally, only the integers 1 through 27 will have to be entered.

The following options then exist:

F1, Enter/ Edit Data

If you select this option, the cursor automatically moves to the position for subject No. 1, card 1-1-1 and data may be entered or edited. The rank assigned to card 1-1-1 by this subject should be entered and then two options exist for entering the data to the computer's memory. Pressing (RET) or down arrow will move the cursor vertically one position down, while pressing the right arrow key will move the cursor to the right one position. The program is set up to accept either the ranks assigned to all 27 cards by the first subject (moving downward), or the ranks assigned by all subjects for a particular card (moving to the right). In either manner, the program will step through all 27 cards for the specified number of subjects. At the end of a row or column of data, use the appropriate arrow keys to move the cursor to the next position on the data entry screen. Every third row on this screen is highlighted to make it easier to keep track of where you are as you scroll through the data. As you reach the end of the displayed data, the screen will scroll to the next position to allow continuation of data entry. Because the scrolling is rather slow, the system is capable of storing up to 11 key presses. This allows you to continue to enter data and then let the system catch up at the end of a row. Due to the delays involved in entering data across rows, it is recommended that, if you have 20 or more subjects, enter the data down the columns.

For editing purposes, use the arrow keys to move the cursor to any position on the screen and make the appropriate change. Pressing either (RET) or

*** ENTER SUBJECT DATA IN THIS TABLE ***

15 SUBJECTS

F1 SAVE DATA
F2 EDIT/ENTER DATA
F3 PRINT DATA
F4 PROGRAM SETUP
ESC MAIN MENU

CARD		1	2	3	4	5	6	7	8
111	N	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
112	B	2.00	10.00	3.00	5.00	2.00	2.00	2.00	2.00
113	W	3.00	19.00	19.00	8.00	8.00	3.00	3.00	3.00
121	F	4.00	2.00	13.00	6.00	4.00	4.00	10.00	4.00
122	J	12.00	12.00	16.00	11.00	5.00	5.00	11.00	5.00
123	C	13.00	20.00	23.00	13.00	13.00	6.00	12.00	6.00
131	X	18.00	5.00	2.00	18.00	14.00	7.00	19.00	7.00
132	S	19.00	14.00	18.00	19.00	15.00	8.00	20.00	8.00
133	M	14.00	23.00	25.00	20.00	19.00	9.00	21.00	9.00
211	U	5.00	3.00	4.00	2.00	3.00	10.00	4.00	10.00
212	G	6.00	11.00	5.00	4.00	6.00	11.00	5.00	11.00
213	Z	7.00	21.00	20.00	9.00	10.00	12.00	6.00	12.00
221	V	8.00	4.00	12.00	3.00	9.00	13.00	13.00	13.00
222	Q	15.00	13.00	15.00	15.00	11.00	14.00	14.00	14.00
223	ZZ	20.00	22.00	21.00	17.00	16.00	15.00	15.00	15.00
231	K	21.00	7.00	6.00	21.00	20.00	16.00	22.00	16.00
232	E	22.00	16.00	7.00	24.00	23.00	17.00	23.00	17.00
233	R	23.00	24.00	22.00	25.00	24.00	18.00	24.00	18.00

Figure 24. Data Entry Screen

any arrow key will make the change in the computer's memory. You must select the "save data" (F2) option at the end of the data entry/edit session to make the changes on disk.

Four special function keys exist to aid in data entry. These allow quick movement to the top, bottom, left, and right of the data matrix, avoiding excessive scrolling. Typical use of these keys would occur during data entry. Upon reaching the end of a row of data, use the "left" function key to move back to the beginning of the row. However, when entering data across rows for 20 or more subjects, the use of the right and left keys leave several columns in the center of the data matrix inaccessible. In this situation, it would be more efficient to enter the data down the columns.

As you exit from entering data by pressing F1, a data checking algorithm is invoked which will indicate possible errors in the data. This algorithm checks the sum of the columns and is able to indicate which columns have probable errors. Usually, the error can be found by comparing the designated column with the same column on the data entry sheet.

F2, Save Data — When you select this option, the program saves the data currently on the data entry screen to a diskette or hard disk file with either the name you previously specified on the Main Menu screen or a new file which you are able to specify. Specifying a new name will create a new file while not changing the original file. Be sure to always save the data set to prevent accidental loss of data.

F3, Print Data — This allows you to obtain a hard copy of the input data set for easy reference. Since the entire data table, when filled, cannot fit on the screen at one time, this may be a more desirable way to proof a data set.

F4, Program Setup — This option directs the program to proceed to the Program Setup screen in order to continue the normal analysis process. Use this option after all data has been entered and checked for accuracy.

ESC, Escape — This option takes you back to the Main Menu and easily allows you to leave the program, help you get reoriented if you were lost or confused, or make

changes in earlier screens. REMEMBER, though, that data are saved to disk only through the Save Data option on this screen.

PROGRAM SETUP

This screen is presented in Figure 25 and gives you options for the types of analyses to be performed. Enter the number(s) of the analyses which you want performed. There are three main analyses:

1. Prototype correlations and Kendall's Coefficient of Concordance
2. Axiom Testing
3. Scaling Solution

```
                 *******   PROGRAM SETUP   *******

TO RUN ANY OF THESE PROGRAMS OR COMBINATIONS  OF PROGRAMS
   CHOOSE THE CORRESPONDING NUMBER(S) AND PRESS   RETURN

          1   PROTOTYPE CORRELATIONS AND KENDALL'S

          2   GROUP AXIOMS

          3   GROUP SCALE

          4   PROTOTYPE AXIOMS

          5   PROTOTYPE SCALE

          6   INDIVIDUAL AXIOMS

          7   INDIVIDUAL SCALES

        ESC   MAIN MENU

              OPTIONS CHOSEN:
```

Figure 25. Program Setup Screen

The prototype correlations analysis performs a Spearman's rank order correlation on each of the subject's rank order data with the rankings associated with the six possible prototype groups. The pattern of correlations indicates the relative importance a subject places on the three dimensions of

the SWAT definition of workload. The Kendall's Coefficient of Concordance is an index of the degree of agreement among the group of subjects about the ordering of the 27 cards. A high Kendall's indicates substantial agreement about the order and therefore about the relative importance of the three dimensions.

The axiom testing section performs the axiom tests for independence, joint independence, and double cancellation. This is done to check for violations of these axioms which may invalidate the additive model as being a suitable model to use for the conjoint scaling routine. Only a summary of the three axiom tests is automatically displayed, but the complete history of results for all the tests may be viewed or printed by selection of the appropriate option.

Because the scaling solution requires data from both the prototype correlations routine and the axiom testing routine, these functions will by default be performed by the program, if you are working on an initial analysis. The prototype information will then automatically be displayed for the user but the axiom test results will not. The prototype information which is displayed will be put into a data file and used by the program for subsequent analyses. This saving of data from the prototype routine will speed up further analyses which you may require.

The scaling solution produces the scale values that result from the conjoint scaling routine. These values are then used as the workload scores in subsequent analyses of results from the Event Scoring phase.

There are three methods for handling the scale development data: group solutions, prototyped solutions, and individual solutions. The determination of which solution is best is based on study objectives and the results of the correlations, Kendall's Coefficient, and axiom tests. For a more detailed explanation of interpreting these analyses, refer to Section 2.5.2, Prototyping, and Section 2.5.3, Sample Data Analysis. Upon choosing a group solution, the program will automatically include all subjects in further analyses. If you specify prototyped solutions, the program will use the suggested groupings provided in the next screen to be described. A separate solution will be provided for each of the three groups. If individual axioms or scaling solutions are chosen, the program will ask for specific individual subjects you desire to analyze separately. This is done by simply entering the subject number(s) and pressing (RETURN). The program will include only their data in the analyses. If you make a mistake and enter an incorrect subject number, reenter that number and they will not be included (subjects to be included are indicated by highlighting).

Upon pressing (RET), the program will proceed with the analyses previously chosen on this Program Setup screen. Due to the criticality of the prototype correlations section, it will ALWAYS be

performed, even if it was not selected. If no changes have been made, however, the program will not have to recalculate this section.

PROTOTYPE ANALYSIS

This screen is presented in Figure 26 and displays the Spearman's rank correlations for each subject's data with the six "perfect" prototype rankings, the suggested prototype groupings based on these correlations, and the Kendall's Coefficient of Concordance for agreement among the entire group of subjects in this analysis. There are several options for this screen:

F1, Change Prototype	It may be desired to group the subjects differently based on the experimenter's evaluation of the pattern of correlations. To do this, use the up and down arrow keys to move the cursor to the appropriate subject's prototype and then enter the new prototype (T, E, or S). The highlighting will change in accordance with the new prototype. To leave a particular subject out of the analysis, simply change that subject's suggested prototype to an "L." The program will ignore this subject's data. This approach can be used in a case where a subject belongs to a particular prototype group, but these data also contain a large number of axiom test violations, indicating an unacceptable amount of error in the data. In this situation, it may be desirable to exclude this subject from the scaling solution, or attempt to remove the ambiguity in the data through additional information obtained from the subject. This is also very useful in examining the Kendall's Coefficient for a particular prototype group. To do this for the time group, for example, change all the effort and stress subjects' suggested prototypes to "L"s and press F1 again to end the changes. The Kendall's Coefficient will automatically be recalculated and displayed.
F2, Print	Sends the correlation matrix, prototypes, and Kendall's Coefficient to a printer.
F3, Return to Program Setup	This option allows you to return to the Program Setup screen and choose additional or different analyses. This option is used in two situations. First of all, if prototype correlations was the only option previously chosen in the Program Setup, F3 allows you to choose additional analyses based on the results of the prototype correlations and Kendall's Coefficient of Concordance. Second, if, based on the results of the prototype correlations and

* PROTOTYPE ANALYSIS OF EACH SUBJECT'S DATA - *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .7046

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

SUBJECTS 15

SUB. #	TES	TSE	ETS	EST	SET	STE	SUGGESTED PROTOTYPE
1	.71	.58	.94	.89	.50	.42	E
2	.40	.56	.36	.50	.98	.99	S
3	.34	.40	.49	.60	.80	.75	S
4	.60	.51	.90	.91	.65	.55	E
5	.78	.70	.90	.85	.60	.56	E
6	1.00	.96	.60	.43	.30	.43	T
7	.60	.43	1.00	.96	.43	.30	E
8	1.00	.96	.60	.43	.30	.43	T
9	.42	.55	.44	.57	.95	.94	S
10	.64	.59	.65	.61	.47	.46	E
11	.44	.55	.49	.61	.95	.93	S
12	.66	.64	.82	.85	.78	.73	E
13	.48	.62	.40	.52	.95	.98	S
14	.99	.96	.64	.49	.40	.51	T
15	.70	.71	.75	.78	.83	.81	S

F1 CHANGE PROTOTYPE
F2 PRINT
F3 RETURN TO PROGRAM SETUP
F4 GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
ESC MAIN MENU

Figure 26. Prototype Analysis Screen

Kendall's Coefficient, you decide to choose different configurations for axiom testing and scaling solutions, you can cancel any additional selections and choose again. For example, say you originally had chosen prototype correlations, group axioms, and group scale. Now you observe that the subjects are not homogeneous, with a Kendall's Coefficient of .72. You may wish to simply proceed with separating the subject population into prototype groups rather than a single group solution. In either case, the program will not recalculate the correlations, expediting continuation through the program. As mentioned earlier, though, you are able to obtain a Kendall's Coefficient for a particular group.

F4, Go to Next Option Chosen in Program Setup	This option will display the results of the next analysis previously chosen, either axiom tests or scaling solution. If no other options were selected, control will be taken back to the Main Menu.
ESC, Escape	This option returns you to the Main Menu in case you run into problems and need to start over.

SUMMARY OF AXIOM TEST VIOLATIONS

When option F4 of the previous screen (Go to Next Option Chosen in Program Setup) is selected, and you had previously chosen axiom tests as an option in the Program Setup screen, the screen shown in Figure 27 will be displayed. This screen presents a summary of the violations for the simple independence axiom tests, which are the most critical axioms for acceptance of an additive model, as well as a summary of the violations of double cancellation and joint independence. The screen heading will state whether the summary is for a single group, a prototype group, or an individual subject. If the summary is for the entire group of subjects, the following options exist:

F1, Go to Next Option Chosen in Program Setup	This option will calculate and display the scaling solution for the group of subjects if scaling was chosen in the Program Setup screen, or whichever other option was selected in the Program Setup screen. If no other options were selected then the program will go back to the Main Menu.
F2, Print Summary of Group Axioms	This will print a hard copy of just this summary of the axiom test violations.
F3, Print Complete Axiom History	This will print a hard copy of the results of the entire set of axiom tests, which includes more detail on the results of the axiom tests. Typically, this

```
******* SUMMARY OF AXIOM VIOLATIONS *******

                    GROUP ANALYSIS

          INDEPENDENCE
T INDEPENDENT OF E AND S =   0. FAILURES OUT OF 108  TESTS
E INDEPENDENT OF T AND S =   0. FAILURES OUT OF 108  TESTS
S INDEPENDENT OF T AND E =   0  FAILURES OUT OF 108  TESTS
          DOUBLE CANCELLATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF   1  TESTS
DOUBLE CANCELLATION IN E x S =   1. FAILURES OUT OF   3  TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   1  TESTS
          JOINT INDEPENDENCE
T x E INDEPENDENT OF S =   4. FAILURES OUT OF 108  TESTS
E x S INDEPENDENT OF T =  11. FAILURES OUT OF 108  TESTS
S x T INDEPENDENT OF E =   4. FAILURES OUT OF 108  TESTS


                  OPTIONS - GROUP
         F1   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
         F2   PRINT SUMMARY OF AXIOM VIOLATIONS
         F3   PRINT COMPLETE AXIOM HISTORY
         ESC  MAIN MENU
```

Figure 27. Axiom Tests Summary Screen

level of detail is not necessary for evaluation of adequacy of an additive model, as the summary of the axiom tests provides the critical information.

WARNING: This printout is at least six pages long, and once F3 is selected, there is no way to abort the print. Be sure you want the entire history printed before selecting F3.

ESC, Escape — As before, this retains the entered data and returns you to the Main Menu in case of problems or for a fast way to exit the program.

If the Axiom Test Summary is for a prototype group, the additional option "Go To Next Prototype" exists, which will display summary information of axiom violations for the next prototype group. Also, selection of either of the print options will print axiom test information for all three prototype groups. If prototype scale was selected on the Program Setup screen, selection of the

"Go to Next Option Chosen in Program Setup" option in this screen will calculate and display scaling information for the first prototype group.

If the summary is for an individual subject, the additional option "Go To Next Individual" exists, which will display this summary information for the next individual selected. Also, either of the print options will print information for each subject chosen. Therefore, if numerous subjects have been included, printing the complete axiom history will result in quite a large amount of information being printed. All other options will operate as previously described.

SCALING INFORMATION

In all three cases, when "Go to Next Option Chosen in Program Setup" is selected on the Summary of Axiom Violations screen and a scaling solution option had been chosen on the Program Setup screen or upon selection of any of the "Scaling" options on the Program Setup screen, the program will display (as shown in Figure 28) the following information on the Scaling Information screen:

1. The last five iterations of the scaling algorithm. For a more detailed explanation of these values, refer to Section 2.5.3, Sample Data Analysis.

2. The rescaled values for each level of the subscales. These are the additive values which, through all possible combinations, form the 27 values of the scaling solution.

3. The approximate relative importance of each factor. This indicates the amount of change from level 1 of a dimension to level 3 of the same dimension.

This information will be from the program-selected algorithm. Refer to Section 2.3.5 for more information on the selection of the appropriate scaling algorithm.

The following options exist on this screen:

F1, Plot of Rescaled Versus Raw Data	This plot, which is depicted in Figure 29, gives an indication of the goodness-of-fit of the rescaled values. The appropriate plot is linearly decreasing from left to right, and data points not lying on the line indicate cards which were displaced from the pattern. While viewing the plot you may either print it out or return to this screen.

```
*******  SCALING INFORMATION  *******
                GROUP SCALE
 LAST 5 ITERATIONS                   THE SCALE VALUES FOR THE ITERATIONS BELOW
ITERATION   THETA    TAU              ARE PRINTED FROM ITERATION NO.    12
    8      .01954  .89744
    9      .01956  .91453                       ADDITIVE     ADDITIVE
   10      .01957  .89744             VARIABLE    MODEL      RESCALED
   11      .01957  .91453             1    TIME 1    -.36        1.55
   12      .01957  .89744             2    TIME 2    -.04       14.46
                                      3    TIME 3     .41       32.72
                                      4 EFFORT 1     -.40         .10
APPROXIMATE RELATIVE IMPORTANCE       5 EFFORT 2      .00       16.21
       OF EACH FACTOR                 6 EFFORT 3      .40       32.42
                                      7 STRESS 1     -.44       -1.65
 31.17  % FOR FACTOR T                8 STRESS 2     -.02       15.52
 32.32  % FOR FACTOR E                9 STRESS 3      .46       34.86
 36.52  % FOR FACTOR S

                         OPTIONS
             F1   PLOT OF RESCALED VS. RAW DATA
             F2   PRINT SCALING INFORMATION
             F3   PRINT ALL ITERATIONS
             F4   VIEW SCALING SOLUTION
             F5   GO TO NEXT OPTION CHOSEN IN PROGRAM SETUP
            ESC   MAIN MENU
```

Figure 28. Scaling Information Screen

```
         PLOT OF ORIGINAL DATA (X-AXIS) VS. DISPARITIES (Y-AXIS)
         *.****.****.****.****.****.****.****.****.
  1.38  .                                         .    1.38
  1.24  *  0                                      *    1.24   F1  PRINT PLOT
  1.10  *                                         *    1.10   F2  RETURN TO
   .96 ..                                         ..    .96       SCALING
   .81  *   0 00                                  *     .81       INFORMATION
   .67  *                                         *     .67   ESC MAIN MENU
   .53  .          0 0                            .     .53
   .39  *       0 0   +                           *     .39
   .24  *                                         *     .24
   .10 ..              0                          ..    .10
  -.04  *               00 00 00                  *    -.04
  -.18  *                                         *    -.18
  -.32  .                       0  0              .    -.32
  -.47  *                         0  00 0         *    -.47
  -.61  *                                         *    -.61
  -.75 ..                                 00      ..   -.75
  -.89  *                                   0     *    -.89
 -1.04  *                                         *   -1.04
 -1.18  .                                    0    .   -1.18
 -1.32  *                                         *   -1.32
        *.****.****.****.****.****.****.****.****.
         . -23.7244. -16.6731.  -9.6218.  -2.5705.   4.4808.  11.5320.
   -27.2500  -20.1987  -13.1474   -6.0962      .9551     8.0064   15.0577
```

Figure 29. Plot of Rescaled Versus Raw Data

F2, Print Scaling Information	Prints a hard copy of the information on this screen as well as the scaling solution for this particular group or individual. This is the typical method of saving a record of the analysis for future reference.
F3, Print All Iterations	Will send the entire listing of parameters, scaling history, original and rescaled values, and scaling solution for both algorithms to a printer. If this analysis is for prototype groups or an individual, this option will print this information for all prototype groups or all selected individuals. Therefore, this option can produce quite a large amount of information; much of which is not necessary for the typical study.
F4, View Scaling Solution	This option will display the scaling solution for this particular group or individual, as presented in Figure 30. Since all 27 combinations cannot fit on the screen at the same time, you can use F2 to display the rest of the combinations. While viewing the scaling solution, an F1 will return you to this screen.

```
                    SCALING SOLUTION
                                                   F1  RETURN TO MENU
STIM   LEVELS   STANDARD   RESCALED                F2  VIEW REST OF
       T E S                                         SCALING SOLUTION
1      1 1 1    -1.200         .0
2      1 1 2     -.777       17.2
3      1 1 3     -.301       36.5
4      1 2 1     -.804       16.1
5      1 2 2     -.380       33.3
6      1 2 3      .096       52.6
7      1 3 1     -.404       32.3
8      1 3 2      .019       49.5
9      1 3 3      .495       68.8
10     2 1 1     -.883       12.9
11     2 1 2     -.459       30.1
12     2 1 3      .017       49.4
13     2 2 1     -.486       29.0
14     2 2 2     -.062       46.2
15     2 2 3      .414       65.5
16     2 3 1     -.086       45.2
17     2 3 2      .337       62.4
18     2 3 3      .813       81.7
19     3 1 1     -.433       31.2
20     3 1 2     -.010       48.3
```

Figure 30. Scaling Solution Screen

F5, Go to Next Option Chosen Program Setup	Continues with any additional analyses which have been chosen in the program setup screen. If no additional options were selected, this option will return you to the Main Menu screen.
ESC, Escape	Returns you to the Main Menu screen.

Appendix G

EVENT SCORING INFORMATION

REFRESHER SUMMARY

This summary outlines the purpose of the Subjective Workload Assessment Technique (SWAT) and the procedure for giving SWAT ratings. SWAT is a quantitative method for measuring mental workload using subjective ratings. Remember that this method uses the ranking information which you provided with your card sort to create a workload scale. This scale has a distinct workload value for each possible combination of Time Load, Mental Effort Load, and Psychological Stress Load, the three dimensions which comprise SWAT. The definitions of the dimensions are as follows: Time Load refers to the amount of task interruption or overlap; Mental Effort Load is the amount of attention or concentration required to perform a task; and Psychological Stress Load refers to the degree of confusion, anxiety, or frustration involved in performing a task. As you may recall, each dimension has three levels of verbal descriptors: low, moderate, and high. The levels of each dimension can be thought of as a three-point scale, with level 1 being the lowest point or least amount of a dimension and level 3 being the highest point or greatest amount of a dimension. Remember that as you give your ratings for a task, do so in the order of Time Load, then Mental Effort Load, and finally Psychological Stress Load. For example, if you decide on a workload rating of 1-2-3 for a trial which you just completed, this indicates that there was enough or extra time to complete the task (1), the task required moderate mental effort (2), and you experienced a high degree of stress (3).

DO YOU HAVE ANY QUESTIONS ABOUT SWAT?

<u>SWAT Dimensions</u>

I. TIME LOAD

1. Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

2. Occasionally have spare time. Interruptions or overlap among activities occur frequently.

3. Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

II. MENTAL EFFORT LOAD

1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

3. Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

III. PSYCHOLOGICAL STRESS LOAD

1. Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

2. Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

3. High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

BIBLIOGRAPHY

Acton, W. H., and Colle, H. A., 1984, The effect of task type and stimulus pacing rate on subjective mental workload ratings, in Proceedings of the IEEE 1984 National Aerospace and Electronics Conference, pp. 818-823, Dayton, Ohio: IEEE.

Acton, W. H. and Crabtree, M. S., 1985, Work' . ı assessment techniques in system redesign, in Proceedings of the IEEE National Aerospace ana ...lectronics Conference, Dayton, Ohio: IEEE.

Albery, W. B., Ward, S. L., and Gill, R. T., 1985, The effect of acceleration stress on human workload, Technical Report AAMRL-TR-85-039, Wright-Patterson Air Force Base, Ohio: Armstrong Aerospace Medical Research Laboratory.

Albery, W. B., Repperger, D. W., Reid, G. B., Goodyear, C., Ramirez, L. E., and Roe, M. M., 1987, Effect of noise on a dual task: Subjective and objective workload correlates, in Proceedings of the IEEE National Aerospace and Electronics Conference, Dayton, Ohio: IEEE.

Arbak, C. J., Shew, R. L., and Simons, J. C., 1984, The use of reflective SWAT for workload assessment, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 959-963, Santa Monica, California: Human Factors Society.

Bates, C., Jr. and Vikmanis, M. M., 1985, Some quantitative methodology for cockpit design, in Proceedings of the Advisory Group for Aerospace Research and Development (AGARD) "Guidance-Control-Navigation Automation for Night All-Weather Tactical Operations" (AGARD-CP-387), pp. 9-1/9-10.

Beare, A. N. and Dorris, R. E., 1984, The effects of supervisor experience and the presence of a shift technical advisor on the performance of two-man crews in a nuclear power plant simulator, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 242-246, Santa Monica, California: Human Factors Society.

Beville Engineering, Inc., 1986, Human factors analysis of refinery consolidation.

Boyd, S. P., The use of conjoint analysis for interval subjective scaling of mental workload, Unpublished Masters Thesis, Virginia Polytechnic Institute, Blackburg, Virginia.

Boyd, S. P., 1983, Assessing the validity of SWAT as a workload measurement instrument, in Proceedings of the Human Factors Society 27th Annual Meeting, pp. 124-128, Santa Monica, California: Human Factors Society.

Courtright, J. F., and Kuperman, G., 1984, Use of Swat in USAF system T&E, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 700-703, Santa Monica, California: Human Factors Society.

Crabtree, M. S., Bateman, R. P., and Acton, W. H., 1984, Benefits of using objective and subjective workload measures, in Proceedings of the Human Factors Society 28th annual Meeting, pp.950-953, Santa Monica, California: Human Factors Society.

Detro, S. D., 1985, Subjective assessment of pilot workload in the advanced fighter cockpit, in Proceedings of the Third Symposium on Aviation Psychology, Columbus, Ohio: The Ohio State University Aviation Psychology Laboratory.

Dodge, D. C., Wong, T. J., and Brown, K. W., 1984, Boom control system improvement study - Phase II: Supplemental indication system, Report No. MDC J9732, Douglas Aircraft Company, McDonnell Douglas.

Eggemeier, F. T., Considerations in the application of subjective measures of workload, in Proceedings of the Ninth Congress of the International Ergonomics Association, pp. 115-117, Bournemouth, England: Taylor and Francis, 1985.

Eggemeier, F. T. 1984, Workload metrics for system evaluation, in Proceedings of the Defense Research Group Panel VIII Workshop "Applications of System Ergonomics to Weapon System Development" (pp. C/5-C/20), Shrivenham, England.

Eggemeier, F. T., 1987, Properties of workload assessment techniques, in P. Hancock and N. Mishkate (Eds.), Human Mental Workload, (Amsterdam, The Netherlands, North Holland, 1988).

Eggemeier, F. T. and Amell, J. R., 1986, Visual probability monitoring: Effects of display load and signal discriminability, in Proceedings of the Human Factors Society 30th Annual Meeting, p. 63, Santa Monica, California: Human Factors Society.

Eggemeier, F. T., Crabtree, M. S., and LaPointe, P. A., 1983, The effect of delayed report on subjective ratings of mental workload, in Proceedings of the Human Factors Society 27th Annual Meeting, pp. 139-143, Santa Monica, California: Human Factors Society.

Eggemeier, F. T., Crabtree, M. S., Zingg, J. J., Reid, G. B., and Shingledecker, C. A., 1982, Subjective workload assessment in a memory update task, in Proceedings of the Human Factors Society 26th Annual Meeting, pp. 643-647, Santa Monica, California: Human Factors Society.

Eggemeier, F. T., McGhee, J. Z., and Reid, G. B., 1983, The effects of variations in task loading on subjective workload rating scales, in Proceedings of the IEEE 1983 National Aerospace and Electronics Conference, pp. 1099-1105, Dayton, Ohio: IEEE.

Eggemeier, F. T., Melville, B. E., and Crabtree, M. S., 1984, The effect of intervening task performance on subjective workload ratings, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 954-958, Santa Monica, California: Human Factors Society.

Eggemeier, F. T. and Stadler, M. A., 1984, Subjective workload assessment in a spatial memory task, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 680-684, Santa Monica, California: Human Factors Society.

Eggleston, R. G., 1984, A comparison of projected and measured workload ratings using the subjective workload assessment technique (SWAT), in Proceedings of the IEEE 1984 National Aerospace and Electronics Conference, pp. 827-831, Dayton, Ohio: IEEE.

Eggleston, R. G. and Kulwicki, P. V., 1984, Estimating the value of emerging fighter/attack system technologies, in Proceedings of the Defense Research Group Panel VIII Workshop "Applications of System Ergonomics to Weapon System Development," pp. 1-20, Shrivenham, England.

Eggleston, R. G. and Quinn, T. J., 1984, A preliminary evaluation of a projective workload assessment procedure, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 695-699, Santa Monica, California: Human Factors Society.

Haskell, B. E. and Reid, G. B., 1987, The subjective perception of workload in low-time private pilots, Journal of Aviation, Space, and Environmental Medicine, pp. 1-12.

Haworth, L. A., Bivens, C. C., Shively, R. J., and Delgado, D., 1987, Advanced cockpit and control configurations for single pilot helicopter-nap-of-the-earth flight. Paper presented at the American Helicopter Society Forty-Third Annual Forum and Technology Display.

Kuperman, G. G., 1985, Pro-SWAT applied to advanced helicopter crewstation concepts, in Proceedings of the Human Factors Society 29th Annual Meeting, (pp. 398-402), Santa Monica, California: Human Factors Society.

Masline, P. J., 1986, A comparison of the sensitivity of interval scale psychometric techniques in the assessment of subjective mental workload, unpublished masters thesis, the University of Dayton, Dayton, Ohio.

Masline, P. J. and Biers, D. W., 1987, An examination of projective versus post-task subjective workload ratings for three psychometric scaling techniques, in Proceedings of the Human Factors Society 31st Annual Meeting, pp. 77-80, Santa Monica, California: Human Factors Society.

Nataupsky, M. and Abbott, T. S., 1987, Comparison of workload measures on computer-generated primary flight displays, in Proceedings of the Human Factors Society 31st Annual Meeting, pp. 548-552, Santa Monica, California: Human Factors Society.

Notestine, J. C., 1984, Subjective workload assessment in a probability monitoring task and the effect of delayed ratings, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 685-689, Santa Monica, California: Human Factors Society.

Nygren, T. E., 1985, Axiomatic and numeric conjoint measurement: A comparison of three methods of obtaining subjective workload (SWAT) rankings, in Proceedings of the IEEE 1985 National Aerospace and Electronics Conference, pp. 878-883, Dayton, Ohio: IEEE.

O'Donnell, R. D. and Eggemeier, F. T., 1986, Workload assessment methodology, in K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), Handbook of Perception and Human Performance, Volume 2: cognitive Processes and Performance, New York: Wiley Interscience.

Ossard, G., Amalberti, R., and Poyot, G., 1987, Evaluation de la charge de travail du pilote induite par un systeme d'arme guide laser (Ministere de la Defense: Centre d'Etudes et de Recherches de Medecine Aerospatiale, Laboratoire d'Etudes Medicophysiologiques 16/330).

Potter, S. S., 1986, Subjective workload assessment technique (SWAT) subscale sensitivity to variations in task demand and presentation rate, unpublished masters thesis, Wright State University, Dayton, Ohio.

Potter, S. S. and Acton, W. H., 1985, Relative contributions of SWAT dimensions to overall subjective workload ratings, in Proceedings of the Third Symposium on Aviation Psychology, pp. 231-238, Columbus, Ohio: The Ohio State University Aviation Psychology Laboratory.

Reid, G. B., 1982, Subjective workload assessment: A conjoint scaling approach, Annual Scientific Meeting, Aerospace Medical Association, pp. 153-154, Bal Harbour, Florida.

Reid, G. B., 1985, The systematic development of a subjective measure of workload, in Proceedings of the Ninth Congress of the International Ergonomics Association, pp. 109-111, Bournemouth, England.

Reid, G. B., 1985, Current status of the development of the subjective workload assessment technique, in Proceedings of the Human Factors Society 26th Annual Meeting, pp. 220-223, Santa Monica, California: Human Factors Society.

Reid, G. B., Eggemeier, F. T., and Nygren, T. E., 1982, An individual differences approach to SWAT scale development, in Proceedings of the Human Factors Society 26th Annual Meeting, pp. 639-642, Santa Monica, California: Human Factors Society.

Reid, G. B. and Nygren, T. E., 1988, The subjective workload assessment technique: A scaling procedure for measuring mental workload, in P. Hancock and N. Mishkate (Eds.), Human Mental Workload, Amsterdam, The Netherlands, Elsevier.

Reid, G. B., Shingledecker, C. A., and Eggemeier, F. T., 1981, Application of conjoint measurement to workload scale development, in Proceedings of the Human Factors Society 25th Annual Meeting, pp. 522-526, Santa Monica, California: Human Factors Society.

Reid, G. B., Shingledecker, C. A., Hockenberger, R. L., and Quinn, T. J. (1984). A projective application of the subjective workload assessment technique, in Proceedings of the IEEE National Aerospace and Electronics Conference, pp. 824-826, Dayton, Ohio: IEEE.

Reid, G. B., Shingledecker, C. A., Nygren, T. E., and Eggemeier, F. T., 1981, Development of multidimensional subjective measures of workload, in Proceedings of the IEEE International Conference on Cybernetics and Society, pp. 403-406.

Schick, F. V., and Hahn, R. L., 1987, The use of subjective workload assessment technique in a complex flight task (AGARD-AG-282), Advisory Group for Aerospace Research and Development, pp. 37-41.

Schlegel, R. E., and Gilliland, K., 1987, Evaluation of the criterion task set, Unpublished Contract Report, Wright-Patterson Air Force Base, Ohio: Air Force Aerospace Medical Research Laboratory.

Schlegel, R.E., Gilliland, K., and Schlegel, B., 1986, Development of the criterion task set performance data base, in Proceedings of the Human Factors Society 30th Annual Meeting, pp. 58-60, Santa Monica, California: Human Factors Society.

Shingledecker, C. A., 1983, Behavioral and subjective workload metrics for operational environments, in Proceedings of the Advisory Group of Aerospace Research and Development (AGARD) "Sustained Intensive Air Operations: Physiological and Performance Aspects" (AGARD-CP-338), pp. 6-1/6-8.

Skelly, J. J. and Purvis, B. D., 1985, B-52 wartime mission simulation: Scientific precision in workload assessment, in Proceedings of the 1985 Air, pp. 105-109.

Thiessen, M. S., Lay, J. E., and Stern, J. A., 1988, Neuropsychological workload test battery validation study, Unpublished Contract Report, Wright-Patterson Air Force Base, Ohio: Air Force Aerospace Medical Research Laboratory.

Vidulich, M. A. and Tsang, P. S., 1985, Assessing subjective workload assessment: A comparison of SWAT and the NASA-BIPOLAR methods, in Proceedings of the Human Factors Society 29th Annual Meeting, pp. 71-75, Santa Monica, California: Human Factors Society.

Warr, D., 1986, A comparative evaluation of two subjective workload measures: Subjective workload assessment technique and the modified Cooper-Harper scale, unpublished masters thesis, Wright State University, Dayton, Ohio.

Warr, D., Colle, H. A., and Reid, G. B., 1986, A comparative evaluation of two subjective workload measures: SWAT and the modified Cooper-Harper scale. Paper presented at the Symposium on Psychology in the Department of Defense, USAFA, Colorado Springs, Colorado.

REFERENCES

Acton, W. H. and Colle, H. A., 1984, The effect of task type and stimulus pacing rate on subjective mental workload ratings, in Proceedings of the IEEE 1984 National Aerospace and Electronics Conference, pp. 818-823, Dayton, Ohio: IEEE.

Arbak, C. J., Shew, R. L., and Simons, J. C., 1984, The use of reflective SWAT for workload assessment, in Proceedings of the Human Factors Society 28th Annual Meeting pp. 959-963, Santa Monica, California: Human Factors Society.

Eggemeier, F. T., Crabtree, M. S., and LaPointe, P. A., 1983, The effect of delayed report on subjective ratings of mental workload, in Proceedings of the Human Factors Society 27th Annual Meeting, pp.139-143, Santa Monica, California: Human Factors Society.

Eggemeier, F. T., McGhee, J. Z., and Reid, G. B., 1983, The effects of variations in task loading on subjective workload rating scales, in Proceedings of the IEEE 1983 National Aerospace and Electronics Conference, pp. 1099-1105, Dayton, Ohio: IEEE.

Eggemeier, F. T., Melville, B. E., and Crabtree, M. S., 1984, The effect of intervening task performance on subjective workload ratings, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 954-958, Santa Monica, California: Human Factors Society.

Emery, D. R. and Barron, F. H., 1979, Axiomatic and numerical conjoint measurement: An evaluation of diagnostic efficacy, Psychometrika, 44, pp. 19'-210.

Green, P. E. and Rao, V., 1971, Conjoint measurement for quantifying judgmental data, Journal of Marketing Research, 8, pp. 355-363.

Green, P. E. and Srinivasan, V., 1978, Conjoint analysis in consumer research: Issues and outlook, Journal of Consumer Research, 5, pp. 103-123.

Holt, J. O. and Wallsten, T. S., 1974, A user's manual for CONJOINT: A computer program for evaluating certain conjoint-measurement axioms, Technical Report No. 42, University of North Carolina, L. L. Thurstone Psychometric Laboratory.

Johnson, R. M., 1973, Pairwise nonmetric multidimensional scaling. Psychometrika, 38, pp. 11-18.

Krantz, D. H. and Tversky, A., 1971, Conjoint measurement analysis of composition rules in psychology, Psychological Review, 78, pp. 151-169.

Kruskal, J. B., 1965, Analysis of factorial experiments by estimating monotone transformations of the data, Journal of the Royal Statistical Society, Series B, 27, pp. 251-263.

Notestine, J. C., 1984, Subjective workload assessment and effect of delayed ratings in a probability monitoring task, in Proceedings of the Human Factors Society 28th Annual Meeting, pp. 685-690, Santa Monica, California: Human Factors Society.

Nygren, T. E., 1982, Conjoint Measurement and Conjoint Scaling: A Users Guide, Technical Report AFAMRL-TR-82-22, Wright-Patterson Air Force Base, Ohio: Air Force Aerospace Medical Research Laboratory.

Nygren, T. E., 1985, An examination of conditional violations of axioms for additive conjoint measurement, Applied Psychological Measurement, 9, pp. 249-264.

Potter, S. S. and Acton, W. H., 1985, Relative contributions of SWAT dimensions to overall subjective workload ratings, in Proceedings of the Third Symposium of Aviation Psychology, pp. 231-238, Columbus, Ohio: The Ohio State University Aviation Psychology Laboratory.

Reid, G. B., Eggemeier, F. T., and Nygren, T. E., 1982, An individual differences approach to SWAT scale development, in Proceedings of the Human Factors Society 26th Annual Meeting, pp. 639-642, Santa Monica, California: Human Factors Society.

Sheridan, T. B. and Simpson, R. W., 1979, Toward the Definition and Measurement of the Mental Workload of Transport Pilots, (FTL Report R79-4), Cambridge, Massachusetts: Massachusetts Institute of Technology Flight Transportation Laboratory.

Ullrich, J. R. and Cummins, D. E., 1973, PCJM: A program for conjoint measurement analysis of polynomial composition rules, Behavioral Science, 18, pp. 226-227.

☆U.S. Government Printing Office: 1989 -- 748-002/00218